Forecasting indoor pollutants concentrations using Fast Fourier Transform (FFT) and Regime Switching Models

Rachid Ouaret, Anda Ionescu, Olivier Ramalho, Viorel Petrehus, Yves Candau

Forecasting indoor pollutant concentrations using Fast Fourier Transform and Regime Switching Models

Rachid Ouaret[1,2], Anda Ionescu [2], Olivier Ramalho[1],
Viorel Petrehus[3] and Yves Candau[2],

{rachid.ouaret}@cstb.fr

Abstract. This paper considers the possibility that the high-frequency formaldehyde concentration follows a regime switching process. It explores the calibration of models built on Threshold Autoregression (TAR) combined with Fast Fourier Transform (FFT) for short-term forecasting of indoor formaldehyde fluctuation. The methodology uses signal decomposition with filtering of the original raw time series into different FFT components removing some high frequencies. Results from FFT-TAR model could predict HCHO concentration patterns with acceptable accuracy and suggest that there is benefit by taking choice of "cutoff" frequency. The mean absolute percentage error (MAPE) of the best scheme for 10-h (600); one day (1440) and 40h ahead out of sample forecasts, were obtained as follows: 3.54%, 14.36% and 15.64%.

Keywords: Forecasting, Threshold Autoregression (TAR), Fast Fourier Transform (FFT), high-frequency, formaldehyde.

1 Introduction

Due to high levels of indoor pollutants and usually 80% of our time spent indoors, significant health problems (eg. respiratory and cardiovascular) can be caused by exposure to indoor air. Therefore, estimating future values of indoor pollutant concentrations is of particular interest in the field of indoor air quality (IAQ). One of the major contaminant in indoor air is formaldehyde (HCHO). It originates from indoor materials such as particleboard, plywood, and combustion (e.g. tobacco smoke) and photochemistry. The dynamics of indoor formaldehyde concentration in real world environments and its relation to sources and climatic variables are complex and little studied. While the outdoor air research has much experience in long monitoring studies with advanced statistical modeling, the field of indoor air quality lacks monitoring studies of pollutant concentrations with a short time-step. The majority of papers published on air quality forecasting uses low-frequency data, i.e. weekly, monthly, quarterly, annual, etc. However, there are very few studies on shorter time interval environmental time series. Granger [1] argued that high-frequency data pose a new set of forecasting problems, demanding new approaches.

In this paper, we propose a new approach in order to predict the indoor pollutant concentration of formaldehyde. The emission mechanisms of HCHO from materials are complex and depend on several parameters that vary over time like temperature and humidity. The temporal variability of these parameters is most often unknown and their physical models are often difficult to build for indoor air pollutants. In this work we develop a method to forecast short-term formaldehyde indoor concentrations using nonlinear time series models and Fourier transform. The succession of phases with high and low variability and the presence of several mean shifts in HCHO concentration over time can be highlighted by regime switching models. In such a case, the dynamic behavior of HCHO depends on the regime that occurs at different time scales. It seems therefore natural to expect that Regime Switching Models may improve prediction of HCHO indoor concentrations. The forecasting model can also be used to provide valuable feedback for the physical processes. To our knowledge, this is the first study exploring HCHO concentration forecasting by nonlinear stochastic time series. The determinist part of HCHO concentration was modeled by multiple Fast Fourier Transform (FFT) which were incorporated within two types of Regimes Switching Models: Smooth Threshold Auto-Regressive (STAR) and Self Exciting Threshold Autoregressive (SETAR) models. In order to compare their performance, the root mean square error (RMSE) and the mean absolute percentage error (MAPE) were calculated for each forecast.

Different sub-samples of the given historical data were used to compare the different forecasting models. The predicted data were calculated up to 2 days ahead at most. The lack of references about forecasting in the field of indoor air quality is the main reason for us to use a multiple stage and hybrid approach. From that perspective, hybridization of many and varied models using complementary and common approaches and strategies is leading to a kind of robustness. Only two-regime switching models are considered in this study. There are mainly two reasons for this: on the one hand, it is difficult to check the stability of multiple regime parameters requiring several statistical hypothesis and, on the other hand, there is no fully elaborate forecasting methodology for these models.

All the computations involved in the present task were carried out on *R* (`http://www.r-project.org/`). The packages Nonlinear Time Series Analysis (RTISEAN) [2] and Nonlinear time series models with regime switching (**tsDyn**) - with personal very small changes –[4, 3] were freely used; all the other complementary functions and simulations were coded by us.

2 Data, statistical proprieties and methodology

The indoor formaldehyde concentration was monitored every minute in an open-plan office occupied by 6-8 persons over 21 days, i.e. 30,000 observations. The trajectory, the histogram, the periodogram and the autocorrelation function are provided in the following figures (Fig. 1 (A)). The observed mean is 7.85 ppb, the standard deviation equals to 3.64 and the kurtosis is not very high (-0.4). The

fluctuation of the HCHO time series shows a relative stability during the first days and some small oscillations during the two days that followed. After that, the general behavior of the time series changed abruptly and became very variable toward the end of the period. The figure 1(C) presents the histogram and density estimates of the distribution of HCHO concentration which reveal three modes and two gaps around 7 ppb and 10 ppb. The data show an asymmetrical distribution with high density concentrations at the left side and a decreasing density towards the right side. The histogram shape reflects the presence of at least two different processes being "mixed" in the HCHO time dynamics. This indicates that the mean of the process shifted over the measuring period. We observe that the data are not second order stationary; there exists a trend with abrupt changes. This explains the form of the autocorrelation function (ACF) that does not quickly decrease towards zero and the spectral density explodes in many frequencies close to zero (fig 1(B and D)). Thus, according to [5] and [6] definitions, the HCHO concentration seems to exhibit a long memory behavior in the covariance sense. Fig.1 (B) shows the periodogram close to zero of HCHO time series; the frequencies corresponding to dominant periods were chosen to construct the first FFT component which represents the general trend, the following FFT components were identified by the same method. Once significant peaks have been isolated from the raw data and the successive residuals; the next step is to rebuild the HCHO in the time domain. The HCHO time series is a mix of several unknown source emissions that makes the extraction of dominant frequency quite difficult because of the random noise. As shown in the Fig.1(B) (more clearly in Fig (3a)), the noise manifests itself across the entire spectrum, including within high frequencies where the time series should have little energy.

Fig.2 illustrates the steps describing the research methodology used in this study. The process begins first with the identification of the non linear class model which is a TAR model. The next step involved identifying the adequate frequencies cutoff of FFTs components which are obtained successively by subtraction (5 time). The FFTs components were modeled using SETAR and LSTAR models by varying the parameters$m$ and τ; that are described later in this paper (see (5)). In reality, using whether SETAR or LSTAR could change very slightly the accuracy of prediction, thus only the SETAR accuracy is presented. In model selection-prediction, both forecasting accuracy (in test set) and classical indicators - Akaike information and Schwarz information criterion - (in learning test) were used to evaluate the most appropriate models[1].

3 Models

3.1 Threshold models

The underlying idea behind the construction of a TAR model for IAQ is to represent the behavior of HCHO concentration by two separate regimes. Generally,

[1] In diagnostic checking tests, we only consider (for now) the accuracy is important, even some models may exhibit the presence of unit root in both regimes.

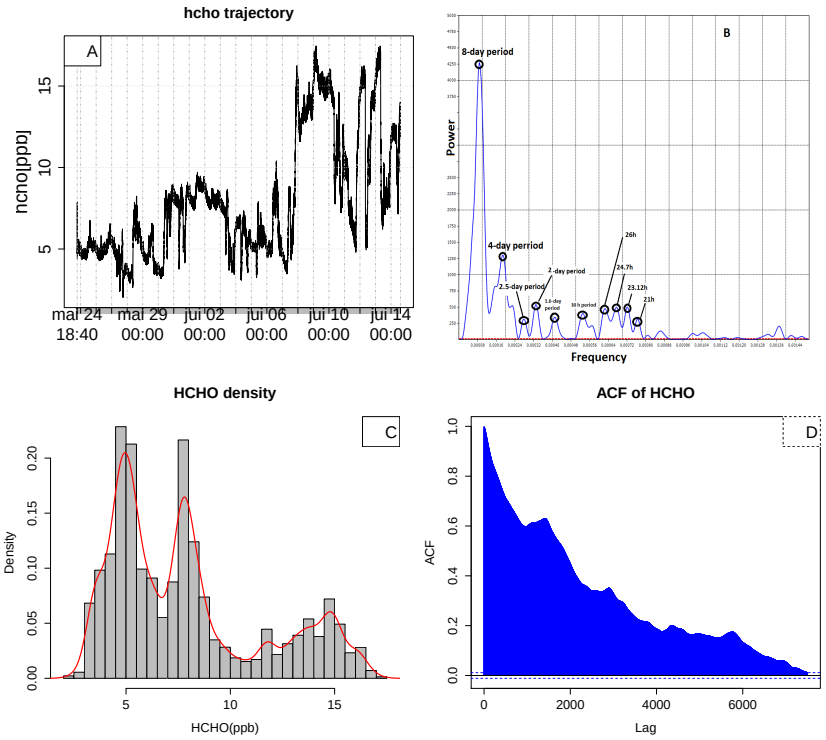


Fig. 1: Descriptive statistics of HCHO: trajectory, raw periodogram, density and ACF.

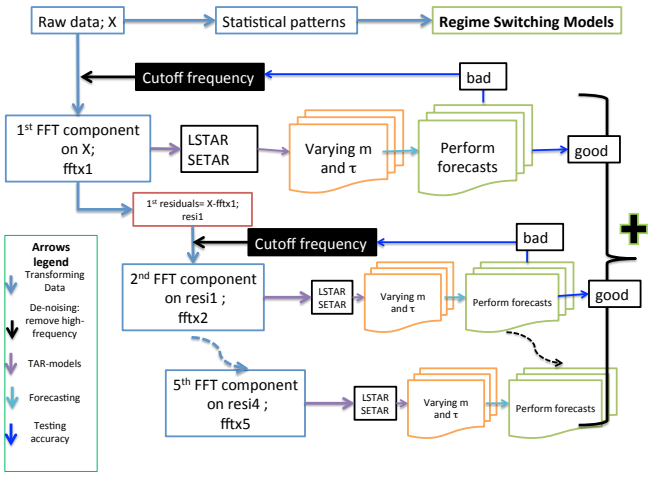


Fig. 2: Research methodology for constructing FFT-TAR model used in the study.

the transitions between regimes in Threshold models (TAR) models are ruled by mechanisms performing with a transition variable, a threshold value and a function of transition. In this study, we restrict our attention to models for which the transition mechanism between regimes is assumed to be observed and, in each regime the dynamics behavior is described by an autoregressive model with an identical delay. Initially proposed extensively by Tong ([7, 9, 8]), the TAR model is defined as follows: for a k -dimensional time series $\{X_t\}_{t\in\mathbb{Z}}$ and for all t, J_t is the transition variable relative to threshold value c; the pair (X_t, J_t) is an TAR with two regimes $TAR(2; p, p)$

$$\begin{array}{c} X_t = \left(\Phi_0^1 + \Phi_1^1 X_{t-1} + \cdots + \Phi_p^1 X_{t-p}\right)\left(1 - \mathbb{I}_{J_t>c}\right) \\ + \left(\Phi_0^2 + \Phi_1^2 X_{t-1} + \cdots + \Phi_p^2 X_{t-p}\right)\left(\mathbb{I}_{J_t>c}\right) + \varepsilon_t \end{array} \tag{1}$$

where $c \in \mathbb{R}$ and ε_t is an i.i.d white noise process conditional upon the $\Omega_{t-1} = \{X_{t-1}, X_{t-2}, \ldots, , X_{t-p}\}$ with $\mathbb{E}\left(\varepsilon_t \mid \Omega_{t-1}\right) = 0$ and $\mathbb{E}\left(\varepsilon_t^2 \mid \Omega_{t-1}\right) = \sigma^2$. The transition function is given by:

$$\mathbb{I}_{J_t>c} = \begin{cases} 1 \;\; is \; J_t > c \\ 0 \; otherwise \end{cases}. \tag{2}$$

A special case arises when the transition variable J_t is taken as the lagged value of the time series itself, that is, it should be any of $\{X_{t-1}, X_{t-2}, \ldots, , X_{t-p}\}$. The resulting model for any lagged value of X_t taken as J_t in equation (1) is Self-Exciting TAR (SETAR). Because of the discrete nature of the threshold variable, the transition between regimes is carried out abruptly. In fact, there are some periods where the time series does not transit abruptly between regimes; therefore it is possible to replace the transition function in (2) by a continuous function $G\left(J_t; \gamma, c\right)$ bounded between 0 and 1; the resulting model is called Smooth Transition AR (STAR) model and is given by:

$$\begin{array}{c} X_t = \left(\Phi_0^1 + \Phi_1^1 X_{t-1} + \cdots + \Phi_p^1 X_{t-p}\right)\left(1 - G\left(J_t; \gamma, c\right)\right) \\ + \left(\Phi_0^2 + \Phi_1^2 X_{t-1} + \cdots + \Phi_p^2 X_{t-p}\right)\left(G\left(J_t; \gamma, c\right)\right) + \varepsilon_t \end{array}, \tag{3}$$

where the parameter γ determines the *smoothness* of the change in the value of the transition function. In this paper, we choose as in many other works ([10]), the logistic function:

$$G\left(J_t; \gamma, c\right) = \frac{1}{1 + \exp\left\{-\gamma\left(J_t - c\right)\right\}} \tag{4}$$

The function $G\left(J_t; \gamma, c\right)$ in the STAR models switches instantaneously between the two regimes as soon as the quantity $(J_t - c)$ changes its sign and this, for γ sufficiently large. Consequently the function $G\left(J_t; \gamma, c\right)$ approximates the indicator function $\mathbb{I}_{J_t}$ in the model TAR (in equation 1) (or SETAR if $J_t = X_{t-d}$ (in equation 3). When $\gamma \to 0$, $(G\left(J_t; \gamma, c\right) \rightsquigarrow 1/2)$ then the STAR model reduces to an AR linear model.

In our work several simulations were carried out to determine γ; the choice of γ is done according to the accuracy of prediction (we have avoided the two extremes mentioned above), and have set it at one ($\gamma = 1$).

In this work, we have chosen to reformulate the models (1) and (3) by introducing the delay time series of *embedding dimension* m and *delay time* τ, and the threshold delay δ in both equations system.

$$\begin{aligned} X_t = &\left(\Phi_0^1 + \Phi_1^1 X_{t-\tau} + \cdots + \Phi_{(m-1)}^1 X_{t-(m-1)\tau}\right)\left(1 - G\left(J_t; \gamma, X_{t-\delta\tau}\right)\right) \\ &+ \left(\Phi_0^2 + \Phi_1^2 X_{t-\tau} + \cdots + \Phi_{(m-1)}^2 X_{t-(m-1)\tau}\right)\left(G\left(J_t; \gamma, X_{t-\delta\tau}\right)\right) + \varepsilon_t \end{aligned} \tag{5}$$

All simulations were carried out by varying m, τ and δ for de-noising FFT components and the choice was done once the models gave us a good forecast. As we can see, doing all possible combinations is not an easy task due to the large data set, because it requires a lot of computational resources. As the frequency of the time series increases the required computational time for simulation become high, demanding a good combination in the models parameters that will make the most efficient best setup.

3.2 Removing noise and time series reconstruction using FFT

The first motivation for working in the frequency domain is that removing noise from a temporal signal can be accomplished, at least approximately, by removing the signal's "high frequency components". The second reason is that many operations are easier, both computationally and conceptually, in the frequency domain. In general way, the harmonic analysis decompose time series over oscillatory waveforms that reveal many properties and provide a path to sparse representations.

FFT is an efficient method for computing the discrete Fourier transform (DFT). The algorithm for computing the DFT was primarily attributed to Cooley and Tukey [12]. Computationally, the DFT is of the order of $\mathcal{O}\left(T^2\right)$ while the FFT is of the order of $\mathcal{O}\left(T \log_2 T\right)$. If $\{x_t\}_{t=0;\ldots,(T-1)}$ is the time series of HCHO, its DFT can be written as:

$$\mathcal{H}_k = \sum_{t=0}^{T-1} x_t e^{-2\pi j \frac{kt}{T}}; \;\; with\; k = 0, \ldots, T-1 \tag{6}$$

As mentioned in the section (2) the random noise pollutes the signal. Thus a very simple approach to noise reduction is to determine the "cutoff" frequencies; this was performed by observing the different resolution after reconstruction i.e based in visual insight. This is done by the inverse FFT (IFFT)

$$x_t = \frac{1}{T} \sum_{t=0}^{T-1} \mathcal{H}_k e^{2\pi j \frac{kt}{T}} \tag{7}$$

Once high frequencies removed from the spectrum, the associated signals can be reconstructed in the time domain using IFFT. Each reconstructed FFT component into the time domain corresponds to the real part of the HCHO Fourier term.

In Figure 3 (a-f) we show the reconstruction based on FFT of the HCHO time series; which represents different patterns and characteristics inside the raw time series. The first FFT component presents the general trend that captures the most important information in terms of representativeness. The rest of oscillatory components shows the different inter-day variability. The second FFT component exhibits roughly the same amplitude over 15 days, after that the shape of the oscillations changed near the end of the meausrement duration. The amplitude modulations of the 3rd, 4th and 5th FFT oscillations are similar, but with small differences in variance. Several simulations were carried out in order to determine the various plausible scenarios. In other word, we begin by whether fixing m and τ and varying the cutoff frequency, or vice versa. The cutoff varying frequencies of each FFT and the parameters of SETAR model for modeling the two different sets are shown in Tab 1. The combined models are compared among each other with corresponding data set for multiple step-ahead forecast.

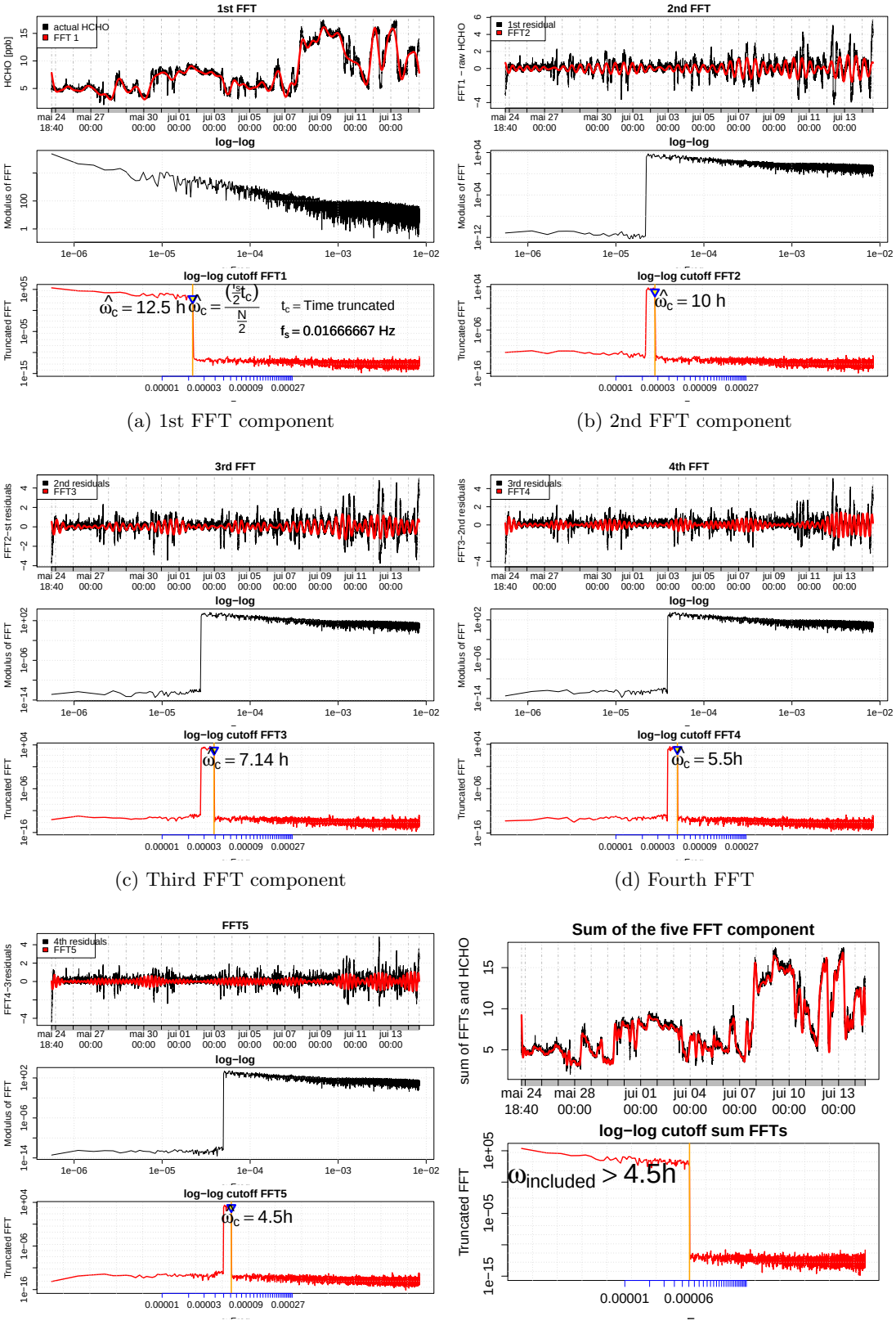
Defining sets	scenarios	Simulation conditions	FFT1	FFT2	FFT3	FFT4	FFT5
			Periods (ω_i) included in the FFTs				
Set1	scenario1	S1; $m = 3, \tau = 1$	2.08 days	25 h	16.6 h	12.5 h	10 h
	scenario2	S2; $m = 3, \tau = 1$	25 h	16.6 h	12.5 h	10 h	7.14 h
	scenario3	S3; $m = 3, \tau = 1$	16.6 h	12.5 h	10 h	8.33 h	6.25 h
	scenario4	S4; $m = 3, \tau = 1$	12.5 h	10 h	8.33 h	7.14 h	5 h
	scenario5	S4s1*	12.5 h	10 h	8.33 h	7.14 h	5 h
	scenario6	S4s2*	12.5 h	10 h	7.14 h	5.5 h	4.5 h
Set2	scenario7	S6s3**	12.5 h	10 h	7.14 h	5.5 h	4.5 h
	scenario8	S5; $m = 3, \tau = 1$	10 h	8.33 h	7.14 h	6.25 h	4.5 h

Table 1: Example of scenarios simulations. * FFT1 was modeled with $m = 2, \tau = 1$ and FFT2 to FFT5 with $m = 3, \tau = 1$. ** FFT2 was modeled with $m = 3, \tau = 1$ and others were modeled with $m = 2, \tau = 1$. For Set 1: learning set=27600; test set=2399; and for Set 2: learning set=15000; test set=2880.

4 Forecasting

The aim of the nonlinear time series considered in our study is to employ the models for forecasting future values of the HCHO variability. Furthermore, out-of sample forecasting was considered as the way to evaluate the capability of regime switching model to capture some of the non-linearity within the data. The forecasting domain has been influenced, for a long time by linear models such as ARMA for which the forecasting function is not very difficult. For example, for AR(2) model, the one step-ahead forecast at time t is $\tilde{x}_{t+1} = \phi_1 x_t + \phi_2 x_{t-1}$ and it not difficult to show that in general it holds that:

$$\tilde{x}_{t+h|t} = \mathbb{E}\left[x_{t+h} \mid \Omega_t\right] \approx \phi_1 \tilde{x}_{t+h-1|t} + \phi_2 \tilde{x}_{t+h|t} \tag{8}$$



(a) 1st FFT component

(b) 2nd FFT component

(c) Third FFT component

(d) Fourth FFT

(e) Fifth FFT component

(f) Sum of the five FFTs.

Fig. 3: Multi-FFT reconstruction of original HCHO time series (scenario 6).

However, considering our case, the time series exhibits nonlinear patterns: asymmetric distribution, abrupt changes, breaks ...etc. Thus forecasting from nonlinear models is a more involved task than forecasting from linear models. That is, the optimal h-step-ahead forecast of x_{t+h} at time t for a general nonlinear model

$$x_t = f\left(x_{t-1}; \theta\right) + \varepsilon_t, \tag{9}$$

is given by:

$$\tilde{x}_{t+h|t} = \mathbb{E}\left[x_{t+h} \mid \Omega_t\right] = \mathbb{E}\left[f\left(x_{t+(h-1)}; \theta\right) \mid \Omega_t\right]. \tag{10}$$

The main difference between the forecast in (10) and the general recursive forecast for an ARIMA (for example) is that the linear expectation operator $\mathbb{E}$ cannot be interchanged with f: the expected value of nonlinear function is not equal to the function evaluated at the value of its arguments [13]. Overall, the relationship between forecasts at different horizons for the nonlinear model does not exist as shown for linear one in the (8). Forecasting 2-step ahead for an nonlinear model is given by

$$\begin{aligned} \tilde{x}_{t+2|t} &= \mathbb{E}\left[f\left(f\left(x_t; \theta\right) + \varepsilon_{t+1}; \theta\right) \mid \Omega_t\right] \\ &= \mathbb{E}\left[f\left(\tilde{x}_{t+1|t} + \varepsilon_{t+1}; \theta\right) \mid \Omega_t\right]. \end{aligned} \tag{11}$$

Unlike the naive approach which considers that $\varepsilon_{t+1} = 0$ in (11) or the possibility of interchanging between $\mathbb{E}$ and f, Monte Carlo simulation and bootstrapping methods are used to compute nonlinear forecasts and to approximate the conditional expectation (11).

The 2-step-ahead Monte Carlo forecast is given by

$$\tilde{x}^{(mc)}_{t+2|t} = \frac{1}{k} \sum_{i=1}^{k} f\left(\tilde{x}_{t+1|t} + \varepsilon_i; \theta\right), \tag{12}$$

where k is some larger number and the ε_t are simulated from prior information about the distribution of ε_{t+1}. While in the bootstrap forecast consider the residuals from the estimated model, $\{\tilde{\varepsilon}_t\}_{t=1,\ldots,n}$ in, the forecast function is given by:

$$\tilde{x}^{(boot)}_{t+2|t} = \frac{1}{k} \sum_{i=1}^{k} f\left(\tilde{x}_{t+1|t} + \tilde{\varepsilon}_i; \theta\right). \tag{13}$$

5 Results and discussion

The forecasting experiment is designed as follows. We use in-sample HCHO data to estimate the parameters of the models of interest for each simulation. Then we make multi out-of-sample forecasts up to 60 minutes, 10 hours, 1 day, 36 hours and 40 hours ahead and evaluate them. We carry out and evaluate forecasts in

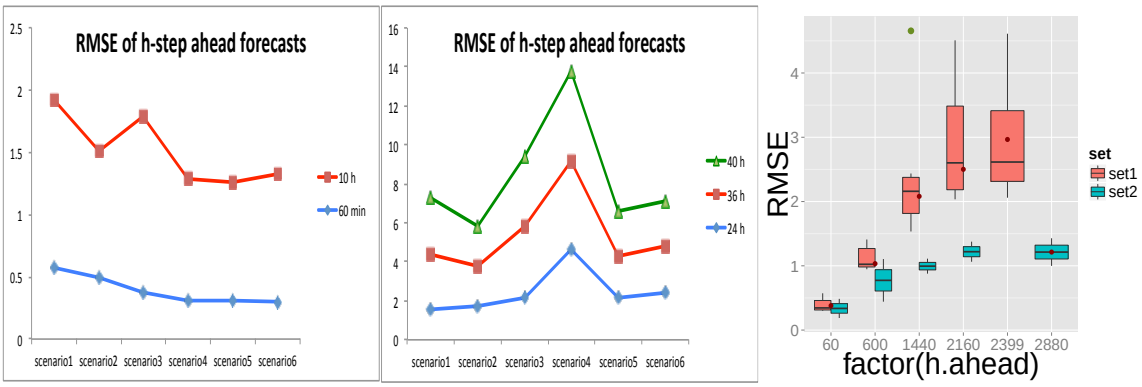


Fig. 4: RMSE of FFT-SETAR (2) for each scenario.

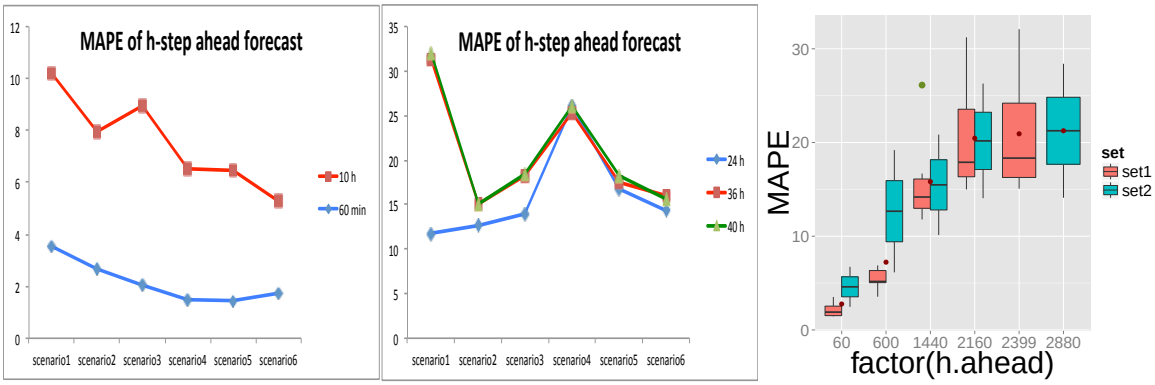


Fig. 5: MAPE of FFT-SETAR (2) for each scenario

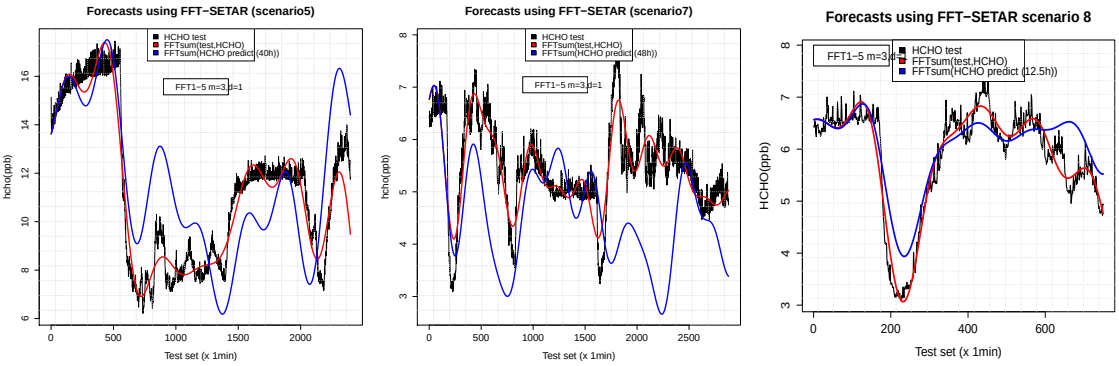


Fig. 6: The indoor HCHO forecating of 40h prediction time horizon from scenarion 5; 48 h forecast from scenario 7 and 12.5 h from scenario 8 (raw data and forecast) (here d is τ).

terms of RMSE and MAPE, we repeat the procedures for each parameter m, τ in the equation (13).

The forecast performance of applied FFT-LSTAR and FFT-SETAR models were evaluated against 60 minutes, 10 hours, 1 day, 36 hours and 40 hours out of sample values; the figures 4 and 5 shows the forecasting performances of the different scenarios models in terms of RMSE and MAPE. For a short-term forecasting 60-600 minutes, the scenarios 4, 5 and 6 outperform the other cases with RMSE around 0.3 ppb and 0.85 ppb respectively. The MAPE values for 60 and 600 out of sample forecasts were below 10%. This shows that the effect of cutoff frequencies in the FFTs have not a significant impact in short-term forecasts. However, if we add more high frequencies in the different FFTs; we reduce consistently the long-term accuracy of forecasting. Specifically, with the same m and τ in scenario 5, the periods included in each FFT are as follows: $> 1\,hour$ (FFT1); $>$ 30 minutes (FFT2); $>$ 15 minutes FFT3; $>$ 7.5 minutes (FF4); $>$ 5 minutes (FFT5). The RMSE values for one day out of sample (1440 min) in this configuration is 3,61 ppb, while it is 2.19 ppb for the scenario 5 (MAPE=28 %; 16% respectively). In general, for the model based on scenario 1, 2, 5, and 6; there is no significant difference with RMSE when the prediction time horizon change between 24 h, 36 h and 40 h.

The models studied in scenario 7 highlights a good performance until 1-day ahead (RMSE =1.1 ppb); and the model obtained from scenario 8 highlights very good accuracy until 12.5 h. The box plot of RMSE illustrates the difference between the prediction in the case of set 1 and set 2 as a function of the forecast horizon; as the prediction time interval increases, the actual HCHO fluctuations deviated from its assumed concentrations, which lead to poor forecasting accuracy. As shown in the (Fig[6]), the HCHO concentrations given by the prediction model of scenario 8 is closely aligned with the raw data. In fact this scheme is the best forecast we got for a 10 hours, with RMSE equal to 0.44 ppb and MAPE equal to 6.14 %. This performance could attributed to the basic underlying characteristic of the good compromise of the included period in FFTs components and the parameters of the SETAR model.

6 Conclusion

This study deals with forecasting indoor pollutant concentrations and in particular HCHO in an open place office, the focus goes beyond the mere estimation and take account of the forecast ability of Regime Switching Models coupled with the fast Fourier transform components. The results of forecast suggest that there is benefit by taking choice of "cutoff" frequency. We have shown that de-noise the raw data permits to identify some characteristics such as beaks. Furthermore, the non-linear models developed here capture the asymmetric behavior of the HCHO fluctuations.

The methodology developed here appears new in order to forecast indoor air quality, the practical implication of the kind of study is that to help decision-

makers to carry out actions that attempt reduce indoor pollution. To do this, it depends on the expected values of indoor pollutants concentrations.

References

1. Granger, C. W. J.: Extracting information from mega-panels and high-frequency data. Statistica Neerlandica, 52: 258–272 (1998)
2. Hegger, R. and Kantz, H. and Schreiber, T.: Practical implementation of nonlinear time series methods: The TISEAN package, CHAOS 9, 413 (1999).
3. Matthieu, S.: tsDyn: Threshold cointegration: overview and implementation in R (2010)
4. Di Narzo, F. Aznarte, J. L. and Matthieu, S.: tsDyn: Time series analysis based on dynamical systems theory (2009)
5. Beran, J.: Statistics for Long-Memory Processes. CRC Press (1994)
6. Guégan, D.: Les chaos en finance: approche statistique. Economica (2003)
7. Tong, H. and Lim, K. S.: Threshold Autoregression, Limit Cycles and Cyclical Data. Journal of the Royal Statistical Society. Series B (Methodological) 42, 3, 245–292 (1980)
8. Tong, H.: Non-linear Time Series: A Dynamical System Approach. Clarendon Press (1993)
9. Tong, H.: Threshold models in non-linear time series analysis. Springer-Verlag (1983)
10. Teräsvirta, T. Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models. Journal of the American Statistical Association 89, 425, 208-218 (1994)
11. Dijk, Dick van and Teräsvirta, T. and Franses, P. H.: Smooth Transition Autoregressive Models-A Survey of Recent Developments. Econometric Reviews 21, 1, 1-47 (2002)
12. Cooley, J.W. and Tukey, J. W.: An Algorithm for the Machine Calculation of Complex Fourier Series. Mathematics of Computation 19, 90 , 297–301 (1965)
13. Franses, P.H. and van Dijk, D.: Non-Linear Time Series Models in Empirical Finance. Cambridge University Press (2000)