www.ietdl.org

Published in IET Signal Processing
Received on 12th February 2014
Revised on 10th June 2014
Accepted on 23rd June 2014
doi: 10.1049/iet-spr.2014.0065

ISSN 1751-9675

Modelling and forecasting of signal-to-interference plus noise ratio in femtocellular networks using logistic smooth threshold autoregressive model

Sepideh Kabiri[1], *Tahereh Lotfollahzadeh*[1], *Mahrokh G. Shayesteh*[1,2], *Hashem Kalbkhani*[1]

[1]*Department of Electrical Engineering, Urmia University, Urmia, Iran*
[2]*Wireless Research Laboratory, ACRI, Electrical Engineering Department, Sharif University of Technology, Tehran, Iran*
E-mail: m.shasyesteh@urmia.ac.ir

Abstract: The aim of this paper is to present a non-linear statistical model to fit and forecast the signal-to-interference plus noise ratio (SINR) in two-tier heterogeneous cellular networks which consist of macrocells and femtocells. Since in these networks the number and locations of femtocell base stations (FBS) are variable, SINR forecasting can be useful in some areas such as power control and handover management. So far, linear autoregressive (AR) models have commonly been used in forecasting the received signal strength (rss) in macrocellular networks. However, AR modelling results in high mean square error (MSE) when data are non-linear. This paper focuses on SINR which takes into account signal strength, interference and noise effects. Moreover, macro-femto cellular network is considered. The *F*-test results show that the SINR data are non-linear, leading to use non-linear models instead of AR model. A non-linear logistic smooth threshold AR (LSTAR) model is utilised to model and forecast the SINR data. Kolmogorov–Smirnov (K-S) test demonstrates that LSTAR provides good fitness to the SINR samples. The results indicate that LSTAR model achieves much better performance in modelling and forecasting of SINR data than the AR model.

1 Introduction

Nowadays mobile devices are widely used two out of three times in indoor environments, commonly home and workplace, and this indoor use is expected to increase in the future years. There is the need for a technology that is able to provide service for indoor environments and deal with the propagation characteristics of inside buildings. An alternative with lower upfront costs is to improve indoor coverage and capacity using the concept of end-consumer installed femtocells or home base stations (BSs) [1]. A femtocell is a low power and short range (up to 30 m) wireless access point that provides in-building coverage to home users and transports the user traffic over the internet-based Internet protocol backhaul such as cable modem or digital subscriber line [2, 3].

Femtocell users experience superior indoor coverage with the wireless networking and can lower their transmit powers. Consequently, femtocells provide higher spatial reuse and cause less interference to other users. The ability to predict the signal-to-interference plus noise ratio (SINR) at the femtocell user equipment (FUE) is necessary to transmit sufficient power from a given femtocell BS (FBS) in order to provide an acceptable quality of service (QoS) over a predetermined service area. Estimating the likely effect of such transmissions on the existing adjacent services is necessary for the improvement of frequency reuse, the implementation of band sharing schemes between different services and for the implementation of cellular systems [4].

SINR is a major challenge in femtocellular networks and greatly affects the QoS. Prediction of rapidly time-varying fading channel conditions enables adaptive data transmission in wireless systems, which in turn improves the QoS for end users and reduces the power consumption for data transmissions. The aim of this paper is to predict the SINR of FUE in two-tier macro-femtocell heterogeneous cellular networks. We focus on the forecasting of time series since most of prediction methods are strictly based on the time series models [5]. We mean from time series the subset of observations obtained from time sequence [6]. Therefore, for prediction, it is necessary to model time series data.

There are some common models used for signal strength forecasting in wireless channels, such as autoregressive (AR) and moving average which are linear models. If data are non-linear, the linear models achieve high mean square errors (MSEs) and cannot provide good performance [7].

Here, we explain several studies on the signal strength prediction. Most of the existing prediction methods in wireless networks use AR models and make the assumption that the input signal is stationary and the channel parameters vary slowly [8]. In [9], Long and Sikdar provided a method for prediction of the non-stationary received signal strength in a more realistic and fast varying wireless environment, using multi-resolution wavelet analysis. They first apply discrete wavelet transform to decompose the strength of the received signal samples into components at different scales, and then use AR and linear

regression models to predict small, medium and large-scale fading components. Finally, they synthesize the output signal of the prediction algorithm. In [10], it is shown that the acceptable performance of AR-based prediction algorithms is achieved because of their lower sampling rates. However, the prediction for more realistic non-stationary data is not improved significantly by lower sampling rates. The iterative AR models used in the method of [10] have the problem of error propagation for prediction steps larger than one. In addition, the network environment is macrocell where path loss (PL) is the impact parameter, whereas femtocell network is not considered in which shadowing plays a major role and consequently rises the complexity.

It is shown that if AR model with a threshold value is considered, the samples smaller than the threshold are modelled with one linear AR process; otherwise there is another AR process. This property makes the model non-linear, where the resulting model is called threshold AR (TAR), proposed in [11, 12]. TAR models have a discontinuous nature as the threshold is passed, and this has led the researchers to consider alternative ways of allowing AR parameters to change, which resulted in the definition of the smooth transition AR (STAR) model [13]. The idea of smooth transition between regimes dates back to [14]. Bacon and Watts [14] both raised the two regression groups, and presented a model, in which the transition from one linear AR to other linear AR is smooth. Hyperbolic function was used for transition. In [13], the logistic function was proposed as transfer function. The logistic STAR (LSTAR) model describes a situation in which the contraction and expansion phases of a time series may have rather different dynamics, and a transition from one regime to other one may be smooth. LSTAR model allows smooth and continuous shift between regimes.

In this paper, to calculate the SINR of FUE located at any distance from its serving BS, we consider the received power from the serving BS, interferences from all co-channel interfering BSs and thermal noise power. Different channel impairments such as PL, shadowing and channel gain are taken into account when calculating the received signal and the interference strengths. Each of the impairments follows a specific distribution, so we expect the received signal and interference strengths and consequently the SINR be non-linear. We demonstrate the non-linearity of SINR data by F-test. Then, we look for non-linear model to fit and forecast the SINR samples. We propose to use non-linear LSTAR model for modelling the SINR of FUE. Using Kolmogorov–Smirnov (K–S) test [15], we demonstrate that the original SINR samples and the samples generated by LSTAR model have the same distribution. Therefore the LSTAR model provides a flexible and appropriate tool for modelling the SINR samples of femtocell users in femtocellular networks. We also aim to forecast five time steps. In the modelling and forecasting steps, the results show the efficiency of LSTAR model in comparison with the AR model.

The rest of this paper is organised as follows. In Section 2, the heterogeneous cellular network model used in this paper is described. In Section 3, we first explain the LSTAR model, next the non-linearity test is introduced to determine whether the data set is non-linear, afterwards the LSTAR model is generated, and then the K–S test is applied to verify the similarity between the distributions of the original SINR data and the generated SINR samples using LSTAR model. In Section 4, simulation results are presented, and finally Section 5 concludes this paper.

2 System model

In the following, the structure of heterogeneous cellular network used in this paper is explained.

The layouts of macrocell network and frequency sub-bands are shown in Fig. 1. Each macrocell has a hexagonal region $\mathcal{H}$ of radius R_{m}. Fractional frequency reuse (FFR) scheme is used in the macrocell network to provide better area spectral efficiency. In the FFR scheme, the area of each macrocell is divided into two non-overlapped regions: inner and edge regions. Macrocell BS (MBS) uses omni-directional antennas to serve the users [15]. It should be noted that in the FFR scheme, the MBS transmission powers of the inner and edge regions users are different where the transmit powers of the edge region users are greater than those of the inner area users, while the powers are fixed in each area.

The available spectrum is partitioned into two parts. The first part is used in the inner region of all macrocells because the frequency reuse factor in this region is one. The second part is utilised in the edge region with the reuse factor equal to three in this paper. Therefore the entire available spectrum is divided into four sub-bands. The radius of the inner region is selected in a manner that the inner and edge regions have the same area [16, 17]. Femtocell network consists of multiple femtocells, each one with radius R_{F}.

It is assumed that the locations of FBSs follow homogeneous spatial Poisson point processes (SPPP) distribution with the intensity λ_{F} [18]. The intensity is considered as the spatial density that shows the average number of FBSs per unit area. The average number of FBSs per cell site is obtained as $N_{\mathrm{F}} = \lambda_{\mathrm{F}}|\mathcal{H}|$ [19], where $|\mathcal{H}|$ denotes the macrocell area. Each FBS uses a fixed power to serve its users. It is assumed that each FBS serves only one user at each time. The algorithm presented in [20] is adopted for channel allocation to the femtocells.

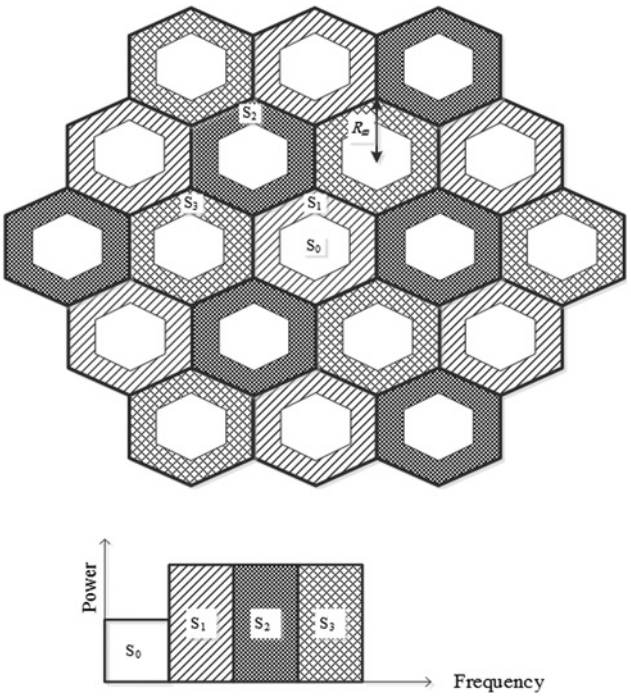

Fig. 1 *Layout of macrocell network and frequency sub-bands used in this paper*

PL, lognormal shadowing and channel gain are used to model the wireless channel. The PLs of different links in the International mobile telecommunications - 2000 (IMT-2000) specification [21] presented in Table 1, are used in this paper. In the Table, f_c denotes the carrier frequency in MHz, W is the wall penetration loss, W^2 is the double-wall penetration loss and D is the distance between the transmitter and receiver.

Shadowing is caused by the obstacles in the propagation between the BS and mobile station (MS) and is characterised by a lognormal distribution as

$$f(\Psi) = \frac{1}{\Psi\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln(\Psi)-\mu)^2}{2\sigma^2}\right) \quad (1)$$

where ln(Ψ) has normal distribution with the mean μ and standard deviation σ. The autocorrelation function for the fluctuation of lognormal shadowing is assumed to be exponential [22] as follows

$$r(\delta d_s) = \sigma^2 \exp\left(-\frac{|\delta d_s|}{d_0}\right) \quad (2)$$

where r denotes the autocorrelation coefficients, δd_s is the distance traversed by the MS and d_0 is the correlation distance (or decay).

The channel gain (H) between the BS and MS is unit mean exponentially distributed as

$$f(H) = \exp(-H) \quad H \geq 0 \quad (3)$$

Therefore the received power (P_R) of the user in the wireless environment, will be equal to [19]

$$P_R = P_T \Psi \phi^{-1} D^{-\alpha} \quad (4)$$

where P_T is the BS transmission power, H denotes the channel gain, Ψ represents the lognormal shadowing with $\mu = 0$ and standard deviation σ in the range 4–10 dB for heterogeneous networks [23], ϕ is the fixed loss, D is the distance between the MS and BS and α is the PL exponent.

The SINR of a FUE positioned at the distance d_F (metres) from the serving FBS is expressed as [19]

$$\text{SINR} = \frac{P_{TF} H_F \Psi_F \phi_F^{-1} d_F^{-\alpha_F}}{N_0 + I_{MF} + I_{FF}} \quad (5)$$

where N_0 is the noise power and I_{MF} is the interference from the macrocell network to the FUE that is calculated as [19]

$$I_{MF} = \sum_{j\in\Omega} P_{TME} H_{MF,j} \Psi_{MF,j} \phi_{MF}^{-1} D_j^{-\alpha_{MF}} \quad (6)$$

Table 1 PLs for different links in IMT-2000 propagation model [21]

Links (T_x/R_x)	Wall loss	Fixed loss	PL exponent	PL
MBS/FUE	W	$\phi_{MF} = W10^{-7.1} f_c^3$	α_{MF}	$\phi_{MF} D^{\alpha_{MF}}$
serving FBS/FUE	0	$\phi_F = 10^{3.7}$	α_F	$\phi_F D^{\alpha_F}$
interfering FBS/FUE	W^2	$\phi_{FF} = W^2 \phi_F$	α_{FF}	$\phi_{FF} D^{\alpha_{FF}}$

where D_j denotes the distance between the FUE and the jth interfering MBS (M_j), $\Omega = \{M_j\}$ represents the set of interfering MBSs and P_{TME} indicates the MBS transmit power of edge region. In (5), I_{FF} is the interference from the co-channel interfering FBSs to the FUE which is computed as [19]

$$I_{FF} = \sum_{i\in\Lambda} P_{TF} H_{FF,i} \Psi_{FF,i} \phi_{FF}^{-1} D_i^{-\alpha_{FF}} \quad (7)$$

where D_i denotes the distance between the FUE and the ith interfering FBS (F_i), $\Lambda = \{F_i\}$ indicates the set of interfering FBSs and P_{TF} represents the FBS transmission power. It is noted that D_is and D_js depend on the locations of FBSs which are generated by SPPP. Therefore the effect of distribution (SPPP) of the locations of the FBSs is reflected in the derivation of SINR through the distances D_js and D_is when computing the interferences in (6) and (7).

3 LSTAR model and its generation for SINR

3.1 LSTAR model

AR is a linear model, while the SINR data are strictly non-linear (will be shown in the next section), so it cannot model the SINR samples properly. TAR model is one of the most important families of non-linear time series models, capable of exhibiting limit cycle behaviour. However, one limitation of this family is that the transitions between various regimes take place in a discontinuous and sudden manner. For a more realistic modelling, these transitions should be smooth; hence the STAR model that moves slightly around the threshold is used. A STAR model allows SINR changes smoothly between the two regimes. Different types of smooth function can be used; in this paper we use the logistic function, where the resulted model is called LSTAR [13].

The LSTAR model of order p for the values of time series ($\boldsymbol{y}_t$), which are the SINR samples in this paper, is defined as [5]

$$\boldsymbol{y}_t = (\phi_1 w_t)(1 - F_{st}(\boldsymbol{y}_{t-d};\, c,\, \gamma)) + (\phi_2 w_t) F_{st}(\boldsymbol{y}_{t-d};\, c,\, \gamma) + u_t \quad (8)$$

where ϕ_1 and ϕ_2 are the parameters of the model with $\phi_j = [\phi_{j0}, \phi_{j1}, \ldots, \phi_{jp}]$ for $j = 1, 2$. $\boldsymbol{w}_t = [1, y_{t-1}, y_{t-2}, \ldots, y_{t-p}]^T$ where $[.]^T$ denotes the transpose of vector, u_t is the zero-mean normally distributed variable, c is a threshold that gives the location of transition, γ is the slope of the transition function, y_{t-d} is the transition variable, d is the delay parameter and F_{st} is the smooth transition function bounded between 0 and 1 that is formulated as [24]

$$F_{st}(y_{t-d};\, \gamma,\, c) = \frac{1}{1 + \exp(-\gamma(y_{t-d} - c))} \quad (9)$$

The main property of LSTAR model is its smooth transition between the two regimes instead of a sudden jump from one regime to another. The transition function $F_{st}(y_{t-d};\, \gamma,\, c)$ is a monotonically increasing function of y_{t-d}. When $\gamma \to \infty$ in (9), if $y_{t-d} < c$, then $F_{st}(y_{t-d};\, c,\, \gamma) = 0$, and if $y_{t-d} > c$, then $F_{st}(y_{t-d};\, c,\, \gamma) = 1$; so (8) becomes a TAR model. When $\gamma \to 0$, the logistic function will be equal to a constant and when $\gamma = 0$, the LSTAR model is reduced to the linear AR model.

doi: 10.1049/iet-spr.2014.0065

In the next sections, the steps of generation of LSTAR model for SINR data in femtocell environments are presented. At first, some tests are applied on the SINR samples to verify that the LSTAR model is valid. Then, the parameters of LSTAR model are obtained. Finally, the SINR forecasting is explained.

3.2 Non-linearity test

In the following, the steps of estimation of the order and the residuals of AR process are explained and then the non-linearity test is described [13].

(1) Estimate AR models of different orders, and select the order of model based on the Akaike's information criterion (AIC) [25]. The AIC criterion is defined as

$$\mathrm{AIC} = n\,\log\left(\sigma_a^2 + 1\right) + 2(p+1) \tag{10}$$

where p and σ_a^2 are the order and noise variance of AR model, respectively, and n is the number of observations. The order that minimises the AIC criterion is selected.
(2) The second step is selecting the value of the delay parameter d. All possible values of $d(1 \le d \le p)$ are examined, and the one that satisfies the following criterion is chosen [24]

$$d = \arg\max_{v \in S} F(p,\ v) \tag{11}$$

where $F(p, v)$ is the F-test statistic of the auxiliary regression with AR(p), v is the delay parameter and S is the set of d values to consider. F-test statistic is calculated as

$$F(p,\ v) = \frac{(\mathrm{SSR}_0 - \mathrm{SSR}_1)/(p+1)}{\mathrm{SSR}_1/(n-p)} \tag{12}$$

where SSR_0 is the square sum of the residuals from the restricted model estimated under the null-hypothesis (model is linear) and SSR_1 denotes the square sum of residual from unrestricted model estimated under the alternative-hypothesis (model is non-linear).
(3) After obtaining the order (p) and the delay value (d), we examine the non-linearity of data by the F-test. The linearity test against non-linearity described in [24] is used here. We should examine the null hypothesis of linearity on (8). We use the auxiliary AR (p) model. Linearity tests against non-linear hypothesis on (8) are

H_0: $\phi_{20} = \phi_{21} = \cdots = \phi_{2p} = 0 \Leftrightarrow \gamma = 0$
H_1: At least one of the $\phi_{2j} \neq 0$ $(j = 1, \ldots, p) \Leftrightarrow \gamma > 0$.

H_0 means that if $\gamma = 0$ the coefficients which make the model non-linear are zero, that is, the model is linear, and H_1 or $\gamma > 0$ implies that at least one of the coefficients is not zero, so the model is non-linear. For non-linearity test, the *pvalue* of F-test statistic is calculated. To calculate the *pvalue*, at first the cumulative distribution function (CDF) of F distribution at F-test statistic is calculated as [13]

$$\mathrm{CDF}\left(F(p,\ d)\right) = \Pr\left(F_{p+1,\,n-p} \le F(p,\ d)\right) \tag{13}$$

where $F_{p+1,\ n-p}$ denotes the F distribution with the $p+1$ and $n-p$ degrees of freedom in the numerator and denominator, respectively. Finally, the *pvalue* of F-test is obtained as

$$pv_{\mathrm{F}} = 2 \times \min\left(CDF\left(F(p,\ d)\right),\ 1 - CDF\left(F(p,\ d)\right)\right) \tag{14}$$

In the F-test, pv_{F} is the probability of obtaining a test statistic at least as extreme as the one that was actually observed, assuming that the null hypothesis is true [26]. One often 'rejects the null hypothesis' when pv_{F} is less than the predetermined significance level which is often set to 0.05 or 0.01. In this paper, if $pv_{\mathrm{F}} < 0.05$ (significance level), the null hypothesis (linearity assumption) is rejected and the non-linearity is accepted. After confirming that the model is non-linear, the next step is to specify the parameters of the LSTAR model.

3.3 Building LSTAR model

After confirming the existence of the LSTAR type non-linearity in SINR time series, we proceed to build a LSTAR model using the ordinary least squares (OLSs) and particle swarm optimisation (PSO).

In statistics, the OLS is a method for estimating the unknown parameters in a linear regression model. This method minimises the sum of squared vertical distances between the observed responses in the data set and the responses predicted by the linear approximation [27].

PSO is a population-based stochastic approach for solving continuous and discrete optimisation problems. In PSO algorithm, each member of the population is called a particle that flies in the multi-dimensional search space with a velocity that is constantly updated by the particle's own experience and the experience of the particle's neighbours or the experience of the whole swarm. The position of a particle represents a candidate solution to the optimisation problem at hand. Each particle searches for the best position in the search space by changing its velocity according to the rules originally inspired by behavioural models of bird flocking [28].

The values of (c, γ) in the LSTAR model ((8)) are found by PSO algorithm, where each pair of $(c,\ \gamma)$ constructs one particle in the two-dimensional search space. The values of $\boldsymbol{\phi}_1 = [\phi_{10},\ \phi_{11},\ \ldots,\ \phi_{1p}]$ and $\boldsymbol{\phi}_2 = [\phi_{20},\ \phi_{21},\ \ldots,\ \phi_{2p}]$ are obtained using the OLS [5]. In the following, the steps of finding the parameters using PSO and OLS algorithms are described.

1. Start with a randomly generated population which are the candidate solutions for c and γ parameters. The parameter c is the threshold that gives the location of the transition function, so its range is between the minimum and maximum amount of SINR data.
2. The following steps are performed for each particle

2.1. The parameters $\boldsymbol{\phi}_1 = [\phi_{10},\ \phi_{11},\ \ldots,\ \phi_{1p}]$ and $\boldsymbol{\phi}_2 = [\phi_{20},\ \phi_{21},\ \ldots,\ \phi_{2p}]$ are denoted by $\boldsymbol{\phi} = \left[\boldsymbol{\phi}_1^{\mathrm{T}},\ \boldsymbol{\phi}_2^{\mathrm{T}}\right]^{\mathrm{T}}$. Note that when c and γ are fixed, the model is linear in the remaining parameters. Estimates of $\boldsymbol{\phi}$ for the kth particle at the ith iteration ($\hat{\boldsymbol{\phi}}_i(k)$) are then easily obtained by OLS as

$$\hat{\boldsymbol{\phi}}_i(k) = \left[\sum_{t=d+1}^{n} \tilde{\boldsymbol{w}}_t(k)\tilde{\boldsymbol{w}}_t^{\mathrm{T}}(k)\right]^{-1}\left[\sum_{t=d+1}^{n} \tilde{\boldsymbol{w}}_t(k)y_t\right] \tag{15}$$

where

$$\tilde{\boldsymbol{w}}_t(k) = \begin{bmatrix} w_t\big(1 - F_{st}\big(y_{t-d};\ c_i(k),\ \gamma_i(k)\big)\big) \\ w_t F_{\mathrm{st}}\big(y_{t-d};\ c_i(k),\ \gamma_i(k)\big) \end{bmatrix} \tag{16}$$

2.2. Using the estimated parameters, the difference between the original and the modelled samples are calculated as follows

$$\varepsilon_t(k) = y_t - \hat{y}_t(k), \quad t = d+1,\ \ldots,\ n \tag{17}$$

where $\hat{\boldsymbol{y}}_t(k) = \tilde{\boldsymbol{w}}_t^{\mathrm{T}}(k)\hat{\boldsymbol{\phi}}(k)$ denotes the modelled samples using the kth particle. The MSE for the kth particle at the ith iteration is computed as

$$\mathrm{MSE}_i(k) = \frac{1}{n-d}\sum_{t=d+1}^{n}\varepsilon_t^2(k) \tag{18}$$

where n is the data length.

2.3. If the MSE (fitness value) of the kth particle, that is, $\mathrm{MSE}_i(k)$, is less than the best fitness value of the kth particle in the history ($pBest(k)$), the current MSE value is set as the new one ($pBest(k)$).

3. The particle with the least MSE value among all the particles is chosen as the $gBest$, that is

$$gBest = \arg\min_{\gamma, c}(\mathrm{MSE}) \tag{19}$$

4. After finding $gBest$, the velocity and positions of each particle are updated as

$$V_{i+1}(k) = V_i(k) + c_1 \times r_1 \times \big(pB(k) - x_i(k)\big) + c_2 \times r_2 \times \big(gBest - x_i(k)\big) \tag{20}$$

$$x_{i+1}(k) = x_i(k) + V_{i+1}(k) \tag{21}$$

where $V_i(k)$ and $x_i(k)$ are the velocity and location (values of pair γ and c) of the kth particle at the ith iteration, respectively. r_1 and r_2 are random numbers in the range [0 1] and c_1 and c_2 are the learning factors in the range [0 2]. Then, the current population is replaced with the new population.

5. If the maximum iteration number or the minimum error criteria are attained, the current values of $\hat{\phi}_i(k_b)$, $\gamma_i(k_b)$ and $\gamma_i(k_b)$, are chosen, where k_b is the particle that gives the minimum MSE in (18); otherwise go back to step 2 and repeat instructions.

For illustration of PSO and OLS steps, a numerical example is presented in Appendix.

3.4 Model verification

The K–S test [29] is used to verify whether the LSTAR model provides a flexible and appropriate tool for modelling the SINR samples of femtocell users. K–S test can be utilised to approve the null hypothesis, that is, the two data sets have the same continuous distribution at a certain level of significance, that is, the two data sets are very similar.

Here, the two-sample K–S test is used. Suppose that data samples, $(z_1,\ \ldots,\ z_{N_1})$ of size N_1 (collected SINR samples), have distribution with the CDF of $Z(z)$ and the second data set samples, $(q_1,\ \ldots,\ q_{N_2})$, of size N_2 (LSTAR data), have distribution with the CDF of $G(z)$. $Z(z)$ and $G(z)$ are the empirical CDFs defined as

$$Z_{N_1}(z) = \frac{1}{N_1}\sum_{i=1}^{N_1}\mathrm{I}(z_i < z) \tag{22}$$

$$G_{N_2}(z) = \frac{1}{N_2}\sum_{i=1}^{N_2}\mathrm{I}(q_i < z) \tag{23}$$

where $I(.)$ is the indicator function that is equal to 1 if $q_i < z$ and is 0, otherwise. Two hypothesises are considered:

H_0: SINR samples follow the LSTAR model, that is, $Z = G$.
H_1: SINR samples do not follow the LSTAR model, that is, $Z \neq G$.

The K–S test based on the KS-statistic, that is, the maximum difference between the two data set samples shown by D_{N_1,N_2} is expressed as follows [15]

$$D_{N_1,N_2} = \sup_z\left|Z_{N_1}(z) - G_{N_2}(z)\right| \tag{24}$$

where the supermom (sup) is the least upper bound of a set S, defined as a quantity β such that no member of the set exceeds β, but if ε is any positive quantity, however small, there is a member that exceeds $\beta - \varepsilon$. The null hypothesis is rejected at level α_{KS} if

$$\sqrt{\frac{N_1N_2}{N_1+N_2}}D_{N_1,N_2} > k_{\alpha_{\mathrm{KS}}} \tag{25}$$

where $k_{\alpha_{\mathrm{KS}}}$ is the critical value that is calculated from the significance level of the test α_{KS} according to the following equation [30]

$$\Pr\left(\sqrt{\frac{N_1N_2}{N_1+N_2}}D_{N_1,N_2} > k_{\alpha_{\mathrm{KS}}}\right) = 1 - \alpha_{\mathrm{KS}} \tag{26}$$

where $\alpha_{\mathrm{KS}} = 0.05$. In this paper, the K–S test is applied to the SINR samples and the LSTAR data where they have the same size, that is, $N_1 = N_2$. In this test, $pvalue$ is the significance level in which the null hypothesis is rejected. $pvalue$ of the K–S (pv_{KS}) test can be approximated using the asymptotic Q-function as

$$pv_{\mathrm{KS}} = \min\left(\max\left(2\times\sum_{k=1}^{101}(-1)^{k-1}\mathrm{e}^{-2(k\xi)^2},\ \ 0\right),\ \ 1\right) \tag{27}$$

where

$$\xi = \max\left(\mathrm{D}_{N_1,N_2}\times\left(\sqrt{\frac{N_1N_2}{N_1+N_2}} + 0.12 + \frac{0.11}{\sqrt{(N_1N_2/N_1+N_2)}}\right),\ \ 0\right) \tag{28}$$

doi: 10.1049/iet-spr.2014.0065

3.5 Forecasting

After fitting the SINR data with the proper non-linear LSTAR model, now it is the time to predict the future values of SINR data.

A forecast might be judged 'successful' if it is close to the outcome, but that judgment may also depend on how 'close' it is measured. One of the key features of measuring forecast accuracy is the division of the data into two separate parts, the first part we refer to as the hold out sample is not used at all during the estimation process and the second part is used for calculating the accuracy of forecasting [31]. It will be used to evaluate the performance after fitting is completed. The procedure will generate a set of five-step-ahead forecasts which will be compared with the hold-out samples.

Linear models such as AR have particular difficulties in forecasting, especially for several steps ahead prediction. The forecasts from non-linear models have to be generated numerically as discussed in [32]. Let $\boldsymbol{y}_t = f(\boldsymbol{y}_{t-1}, \ldots \boldsymbol{y}_{t-p}; r) + u_t$ be a non-linear model, where r is the parameter vector and $u_t$is a zero-mean normally distributed variable. One-step ahead forecasting can be expressed as [32]

$$\hat{y}_{t+1|t} = f\left(y_t,\ \ldots,\ y_{t-p+1};\ \hat{\theta}_t\right) \tag{29}$$

where $\hat{\theta}_t$ indicates that parameters estimates are obtained using the observations up to time period t. Thus, if one knows $f(.)$ or has an acceptable approximation for it, one-step forecast can be achieved with no difficulty. The two-step forecast can be written as follows

$$\hat{y}_{t+2|t} = f\left(\hat{y}_{t+1|t} + u_t,\ y_t,\ \ldots,\ y_{t-p+2};\quad \hat{\theta}_t\right) \tag{30}$$

From the equation, it is clear that each step of forecast is performed based on the previous data set. Our aim is to provide an accurate and low-complexity long-range prediction.

Noting the above, in the LSTAR model, one-step ahead forecast is obtained as

$$\begin{aligned}\hat{y}_{t+1|t} = {} & (\phi_1 w_{t+1})\left(1 - F_{\text{st}}\left(y_{t-d+1};\ \gamma,\ c\right)\right) \\ & + \left(\phi_2 w_{t+1}\right)F_{\text{st}}\left(y_{t-d+1};\ \gamma,\ c\right) + u_{t+1}\end{aligned} \tag{31}$$

And the two-step forecast is expressed as (see (32))
More generally, five-step ahead forecast will be computed recursively as (see (33))

4 Results

In this section, we present simulation results using the system parameters listed in Table 2. The cellular network setup consists of a central hexagonal macrocell surrounded by two rings of interfering macrocells as shown in Fig. 1. The femtocells in the coverage of the central macrocell are analysed. We use SPPP distribution to generate 100 and 200 FBS locations. For each number of FBSs, 10^4 'Monte Carlo' trials were carried out where in each trial we generated 100, 200 and 300 received SINR samples. Each FBS serves only one FUE and the distance of FUE from the serving FBS in the FBS coverage area is randomly chosen. Simulations were performed for noise free and noisy cases. In the noisy case, two different noise powers are considered: −120 and −174 dBm/Hz which are commonly used values. It should be noted that we fit and forecast the SINR in dB scale.

Table 2 Simulation parameters

Symbols	Descriptions	Values
R_{M}, R_{F}	macrocell, femtocell radius	500, 30 m
P_{TMC}, P_{TME}	macrocell transmission power for central and edge regions	40, 43 dBm
P_{TF}	femtocell transmission power	20 dBm
W	wall penetration loss	24 dB
α_{MF}, α_{F}, α_{FF}	PL exponents of different links	3.5, 2, 2
σ_{MF}, σ_{FF}, σ_{F}	standard deviations of shadowing of different links	8, 4, 4 dB

We use the low-pass exponential filter $h_w(k)$ with the length five for smoothing the channel effects which is formulated as [33]

$$h_w(k) = \frac{1}{d_{\text{avg}}}\exp\left(-\frac{k}{d_{\text{avg}}}\right), \quad k \geq 0 \tag{34}$$

where d_{avg} is the effective length of the smoothing window; in this paper we use $d_{\text{avg}} = 5$. The smoothed data are obtained by the convolution of the original SINR data and the low-pass exponential filter. In Fig. 2, the original and smoothed data are shown for 300 received SINR samples.

4.1 Autoregressive model specification

To select the best order (p), we examine different orders and select the one that yields the minimum AIC. In Table 3, we have demonstrated the average AIC values of various orders for 100 and 200 FBSs and 100, 200 and 300 data lengths where 10^4 data records (SINR samples) are generated for each case. We observe that the minimum values of AIC for both number of FBSs are obtained as: $p = 1$ for 100 SINR samples; $p = 2$ for 200 SINR samples; and $p = 3$ for 300 SINR samples. It is seen that when the length of data increases, the order of the autocorrelation model increases as well. Hence, we choose the mentioned orders.

$$\begin{aligned}\hat{y}_{t+2|t} = {} & \left(\phi_{10} + \phi_{11}\hat{y}_{t+1|t} + \cdots + \phi_{1p}y_{t-p+2}\right)\left(1 - F_{\text{st}}(y_{t-d};\ \gamma,\ c)\right) \\ & + \left(\phi_{20} + \phi_{21}\hat{y}_{t+1|t} + \cdots + \phi_{2p}y_{t-p+2}\right)F_{\text{st}}\left(y_{t-d+2};\ \gamma,\ c\right) + u_{t+2}\end{aligned} \tag{32}$$

$$\begin{aligned}\hat{y}_{t+5|t} = {} & \left(\phi_{10} + \phi_{11}\hat{y}_{t+4|t} + \phi_{12}\hat{y}_{t+3|t} + \phi_{13}\hat{y}_{t+2|t} + \phi_{14}\hat{y}_{t+1|t} + \cdots + \phi_{1p}y_{t-p+5}\right)\left(1 - F_{\text{st}}(y_{t-d+5};\ \gamma,\ c)\right) \\ & + \left(\phi_{20} + \phi_{21}\hat{y}_{t+4|t} + \phi_{22}\hat{y}_{t+3|t} + \phi_{23}\hat{y}_{t+2|t} + \phi_{24}\hat{y}_{t+1|t} + \cdots + \phi_{2p}y_{t-p+5}\right)F_{st}\left(y_{t-d+5};\ \gamma,\ c\right) + u_{t+5}\end{aligned} \tag{33}$$

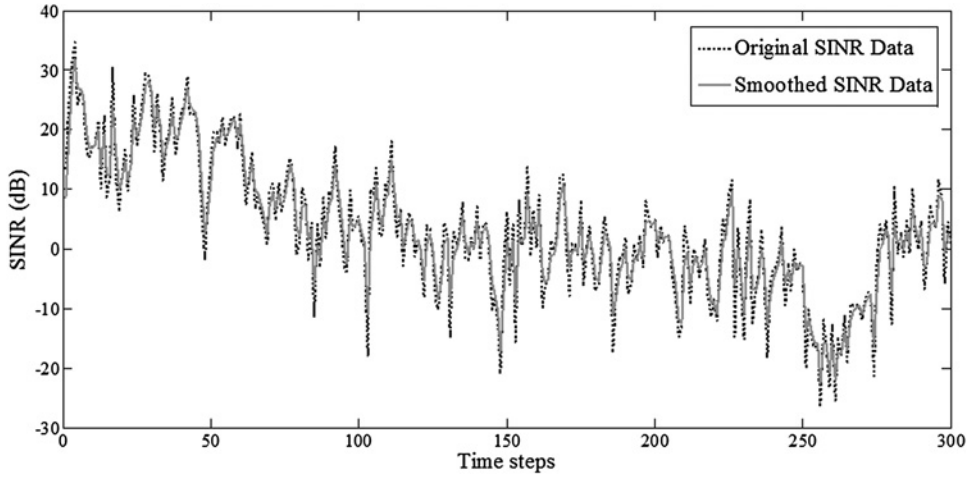

Fig. 2 *Original and smoothed SINR data*

Table 3 Average AIC values

Number of FBSs Order (p)	100			200		
	Data length					
	100	200	300	100	200	300
1	**149.79 (minimum)**	289.67	429.24	148.86 (minimum)	288.71	427.04
2	150.47	**288.42 (minimum)**	425.98	149.50	**287.53 (minimum)**	423.72
3	151.73	288.87	**425.57 (minimum)**	150.76	288.05	**423.37 (minimum)**
4	153.28	290.08	426.41	152.31	289.25	424.20
5	154.91	291.54	427.68	153.94	290.70	425.44
6	156.57	293.09	429.10	155.61	292.24	426.85

4.2 Delay parameter estimation and non-linearity test

After selecting the order (p), we try to find the delay of the series (d). For $1 \le d \le p$, the F-test statistic has been calculated and listed in Table 4. We observe that for the data lengths 100, 200 and 300, $d = 1$ gives the maximum F-test statistic; hence, $d = 1$ is selected as the delay value.

To test the non-linearity, we consider the *pvalue* of F-test. As seen from Table 5, the *pvalue*s for data lengths of 100, 200 and 300 are much smaller than the critical value 0.05. Therefore we reject the null hypothesis that indicates linearity, and accept the non-linearity assumption, that is, SINR data samples are non-linear.

4.3 LSTAR model specification

Finally, the model is estimated and optimised by the OLS and PSO. The average values of γ and c are reported in Table 6. The estimated γ value suggests the transition speed from one regime to the other one.

In Fig. 3, the performances of AR and LSTAR in modelling SINR data for 300 samples are shown. It is clear that LSTAR model fits to the non-linear data very well, whereas AR model fails. Therefore LSTAR is proper for modelling of SINR in macro-femto cellular networks.

The CDFs of MSEs for AR and LSTAR models in data fitting for different number of FBSs are, respectively, presented in Figs. 4 and 5, where noise power is set to −174 dBm/Hz. The range of MSEs in AR model (0–1000) is much greater than that of LSTAR model (0–25), that is, LSTAR modelling yields much less MSE than the AR. The CDFs of AR model for MSE values < 100 are rather the same for all values of data length and FBSs. The same holds true for MSEs < 5 in LSTAR modelling. For MSEs more than 100 in AR model and MSEs > 5 in LSTAR model, the CDFs of higher data lengths become greater.

In Table 7, the average MSEs of AR and LSTAR models are presented for the noise power equal to −174 dBm/Hz.

Table 4 F-test statistic ($F(p, d)$) values for different number of FBSs and data lengths, note that $1 \le d \le p$

Number of FBSs Delay (d)	100			200		
	Data length					
	100	200	300	100	200	300
1	3195.9	2962.2	2812.5	3734.7	3213.3	3096.4
2	–	2948.6	2795.5	–	3198.8	3077.8
3	–	–	2786.2	–	–	3068.4

doi: 10.1049/iet-spr.2014.0065

Table 5 pv_F of F-test statistic for different number of FBSs and data lengths, note that $1 \le d \le p$

Number of FBSs	100			200		
Delay (d)	Data length					
	100	200	300	100	200	300
1	6.8×10^{-3}	421.7×10^{-6}	92.9×10^{-6}	803.1×10^{-6}	271.8×10^{-6}	62.2×10^{-6}
2	–	425.1×10^{-6}	94.4×10^{-6}	–	273.9×10^{-6}	63.1×10^{-6}
3	–	–	95.2×10^{-6}	–	–	63.6×10^{-6}

Table 6 Values of γ and c

Number of FBSs	100			200		
Data length	100	200	300	100	200	300
γ	5.21	5.04	4.83	5.28	5.06	4.82
c	8.37	7.99	8.25	7.21	8.73	9.15

We observe that the LSTAR achieves much better performance than the AR model.

4.4 Model verification

The ratio of the number of accepted K–S tests to the total number of performed K–S tests for different number of FBSs and SINR samples is provided in Table 8. It is observed that the minimum ratio is 86.95%. This means that the null hypothesis is accepted for most of data samples, that is, the collected SINR samples and LSTAR data follow the same distribution. In addition, it is seen that by increasing the number of samples, the number of accepted K–S tests increases, which implies the parameters of the LSTAR model are determined with higher accuracy.

4.5 SINR forecasting

The MSEs of forecasting using AR and LSTAR models for different time steps are depicted in Figs. 6 and 7, respectively, where noise power is −174 dBm/Hz. We observe that for the same number of FBSs and data length, the MSE of the LSTAR model is less than that of the AR model. Hence, for all steps of prediction, LSTAR model is more accurate. The one-step ahead forecast yields the least MSE and five-step ahead gives the highest value of MSE, as expected. The results indicate that the LSTAR model can be used in modelling and forecasting of SINRs of femtocell users in heterogeneous cellular network.

4.6 Noise power effect on the data fitting and forecasting

Now, we analyse the effect of noise power on the modelling and forecasting of SINR samples. In Tables 9 and 10, the average MSEs of AR and LSTAR models in data fitting for different noise powers are, respectively, presented. As shown, because of the little difference between the cases of no noise and −174 dBm/Hz noise power, the same results

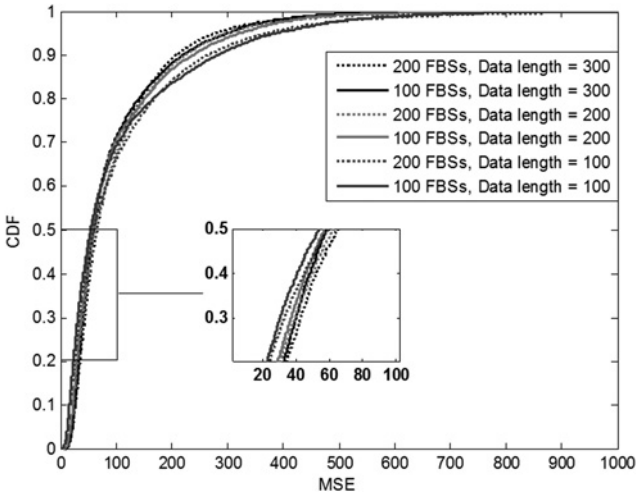

Fig. 4 *CDF of MSE for AR model in SINR data fitting (noise power = −174 dBm/Hz)*

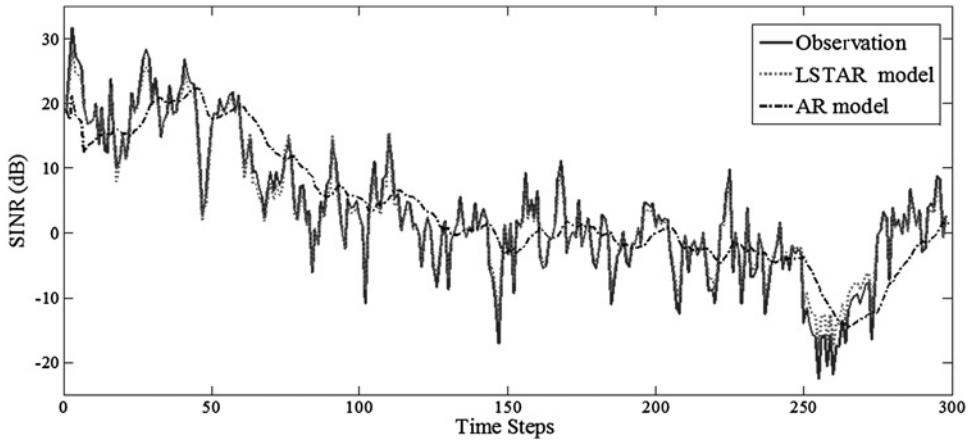

Fig. 3 *SINR modelling with the AR and LSTAR models*

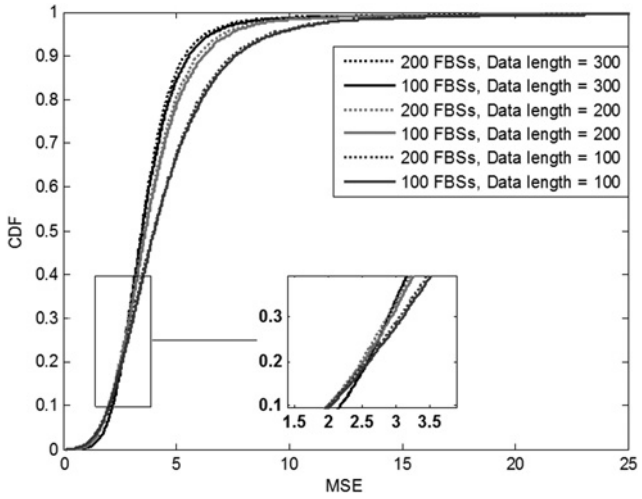

Fig. 5 *CDF of MSE for LSTAR model in SINR data fitting (noise power = −174 dBm/Hz)*

Table 7 Average MSEs of different models in SINR data fitting for noise power equal to −174 dBm/Hz

Number of FBSs	100			200		
Data length	100	200	300	100	200	300
MSE of AR	114.78	104.71	100.82	111.59	101.66	99.41
MSE of LSTAR	4.71	4.12	3.87	4.61	3.98	3.82

Table 8 Results of the K–S test

Number of FBSs	100			200		
Data length	100	200	300	100	200	300
ratio of accepted K–S tests, %	86.95	90.34	90.96	86.78	89.62	90.78

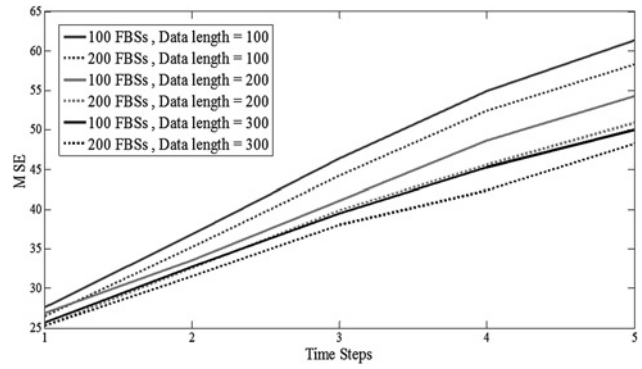

Fig. 6 *MSEs of forecasting of SINR data for AR model (noise power = −174 dBm/Hz)*

Legends are in the same order as the curves

are obtained for the two schemes. When noise power increases to −120 dBm/Hz, the average MSE for LSTAR changes very slightly; this implies that LSTAR model fits very well to the SINR data for different noise powers.

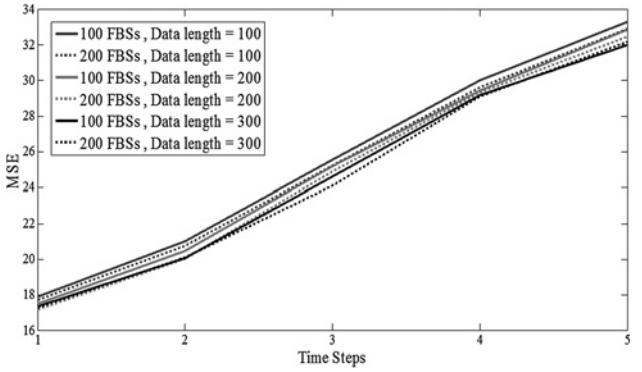

Fig. 7 *MSEs of forecasting of SINR data for LSTAR model (noise power = −174 dBm/Hz)*

Legends are in the same order as the curves

Table 9 Average MSE of AR model in data fitting for different noise powers

Number of FBSs	100			200		
Noise power	Data length					
	100	200	300	100	200	300
0 (no noise)	114.77	104.70	100.82	111.59	101.65	99.40
−174 dBm/Hz	114.77	104.70	100.82	111.59	101.65	99.41
−120 dBm/Hz	104.18	95.91	92.78	107.08	97.85	95.82

Table 10 Average MSE of LSTAR model in data fitting for different noise powers

Number of FBSs	100			200		
Noise power	Data length					
	100	200	300	100	200	300
0 (no noise)	4.75	4.12	3.89	4.61	3.98	3.81
−174 dBm/Hz	4.76	4.13	3.90	4.61	3.96	3.82
−120 dBm/Hz	4.58	3.98	3.81	4.52	3.89	3.75

Comparing Tables 9 and 10, we conclude that LSTAR achieves much less MSE than the AR in all cases.

In Tables 11 and 12, the average MSEs of five time step forecasting for different noise powers are demonstrated for AR and LSTAR models, respectively. Like the data fitting step, for the cases of no noise and −174 dBm/Hz noise

Table 11 Average MSE of AR model in five time step forecasting for different noise powers

Number of FBSs	100			200		
Noise power	Data length					
	100	200	300	100	200	300
0 (no noise)	45.43	40.83	38.65	43.36	38.87	37.09
−174 dBm/Hz	45.43	40.89	38.66	43.36	38.87	37.09
−120 dBm/Hz	43.81	40.77	37.86	42.49	38.26	36.42

doi: 10.1049/iet-spr.2014.0065

Table 12 Average MSE of LSTAR model in five time step forecasting for different noise powers

Number of FBSs	100			200		
Noise power	Data length					
	100	200	300	100	200	300
0 (no noise)	25.50	24.83	24.83	25.63	25.11	24.88
−174 dBm/Hz	25.56	25.09	24.63	25.94	24.78	24.56
−120 dBm/Hz	25.42	24.86	24.70	25.56	25.02	24.85

power, the results are the same, but when the noise power increases, MSE decreases a little. Moreover, the LSTAR model significantly outperforms the AR model.

5 Conclusion

In this paper, we studied the modelling and forecasting of SINR data in femtocellular systems. It was demonstrated that the linearity test for SINR samples of femtocell users was rejected which means the SINR samples cannot be modelled with linear models such as AR. Then, we tried to figure out if it is appropriate to fit the non-linear data with LSTAR model which allows the transition between regimes smoothly rather than a sudden jump. K–S test showed that the CDFs of the SINR samples and the LSTAR model are the same, so LSTAR is a proper model to fit into the SINR data. Based on the LSTAR model, forecasting method was used for prediction of SINRs of femtocell users. The results showed that the performance of LSTAR is much better than the AR in both modelling and forecasting, that is, LSTAR can be used in the telecommunication forecasting issues.

6 References

1 Chandrasekhar, V., Andrews, J., Gatherer, A.: 'Femtocell networks: a survey', *IEEE Commun. Mag.*, 2008, **46**, pp. 59–67
2 Zahir, T., Arshad, K., Ko, Y., Moessner, K.: 'A downlink power control scheme for interference avoidance in femtocells'. 2011 Seventh Int. Wireless Communications and Mobile Computing Conf. (IWCMC), 2011, pp. 1222–1226
3 López-Pérez, D., Valcarce, A., De La Roche, G., Zhang, J.: 'OFDMA femtocells: a roadmap on interference avoidance', *IEEE Commun. Mag.*, 2009, **47**, pp. 41–48
4 Chandrasekhar, V., Kountouris, M., Andrews, J.G.: 'Coverage in multi-antenna two-tier networks', *IEEE Trans. Wirel. Commun.*, 2009, **8**, pp. 5314–5327
5 Franses, P.H., Van Dijk, D.: 'Non-linear time series models in empirical finance' (Cambridge University Press, New York, USA, 2000)
6 Box, G.E., Jenkins, G.M., Reinsel, G.C.: 'Time series analysis: forecasting and control' (John Wiley & Sons, USA, 2013)
7 Chatfield, C.: 'Time-series forecasting' (Chapman and Hall/CRC, USA, 2002)
8 Semmelrodt, S., Kattenbach, R.: 'Investigation of different fading forecast schemes for flat fading radio channels'. 2003 IEEE 58th Vehicular Technology Conf., 2003. VTC 2003-Fall., 2003, pp. 149–153
9 Long, X., Sikdar, B.: 'A wavelet based long range signal strength prediction in wireless networks'. 2008 IEEE Int. Conf. on Communications, ICC'08, 2008, pp. 2043–2047
10 Eyceoz, T., Hu, S., Duel-Hallen, A.: 'Performance analysis of long range prediction for fast fading channels'. Proc. of the 33rd CISS, 1999
11 Tong, H.: 'On a threshold model' (Sijthoff & Noordhoff, 1978)
12 Tong, H., Lim, K.S.: 'Threshold autoregression, limit cycles and cyclical data', *J. R. Stat. Soc.. Ser. B (Methodol.)*, 1980, **42**, pp. 245–292
13 Teräsvirta, T., Tjøstheim, D., Granger, C.: 'Aspects of modelling nonlinear time series', *Handbook of econometrics*, 1994, **4**, pp. 2917–2957
14 Bacon, D.W., Watts, D.G.: 'Estimating the transition between two intersecting straight lines', *Biometrika*, 1971, **58**, pp. 525–534
15 Ekstrom, H., Furuskar, A., Karlsson, J., *et al.*: 'Technical solutions for the 3G long-term evolution', *IEEE Commun. Mag.*, 2006, **44**, pp. 38–45
16 Novlan, T., Andrews, J.G., Sohn, I., Ganti, R.K., Ghosh, A.: 'Comparison of fractional frequency reuse approaches in the OFDMA cellular downlink'. 2010 IEEE Global Telecommunications Conf. (GLOBECOM 2010), 2010, pp. 1–5
17 Lee, P., Lee, T., Jeong, J., Shin, J.: 'Interference management in LTE femtocell systems using fractional frequency reuse'. 2010 12th Int. Conf. on Advanced Communication Technology (ICACT), 2010, pp. 1047–1051
18 Chandrasekhar, V., Andrews, J.: 'Uplink capacity and interference avoidance for two-tier femtocell networks', *IEEE Trans. Wirel. Commun.*, 2009, **8**, pp. 3498–3509
19 Kim, Y., Lee, S., Hong, D.: 'Performance analysis of two-tier femtocell networks with outage constraints', *IEEE Trans. Wirel. Commun.*, 2010, **9**, pp. 2695–2700
20 Lee, J.Y., Bae, S.J., Kwon, Y.M., Chung, M.Y.: 'Interference analysis for femtocell deployment in OFDMA systems based on fractional frequency reuse', *IEEE Commun. Lett.*, 2011, **15**, pp. 425–427
21 Recommendation I.: '1225, 'Guidelines for evaluation of radio transmission technologies for IMT-2000'' (International Telecommunication Union, 1997)
22 Gudmundson, M.: 'Correlation model for shadow fading in mobile radio systems', *Electron. Lett.*, 1991, **27**, pp. 2145–2146
23 Erceg, V., Greenstein, L.J., Tjandra, S.Y., *et al.*: 'An empirically based path loss model for wireless channels in suburban environments', *IEEE J. Sel. Areas Commun.*, 1999, **17**, pp. 1205–1211
24 Luukkonen, R., Saikkonen, P., Teräsvirta, T.: 'Testing linearity against smooth transition autoregressive models', *Biometrika*, 1988, **75**, pp. 491–499
25 Akaike, H.: 'A Bayesian extension of the minimum AIC procedure of autoregressive model fitting', *Biometrika*, 1979, **66**, pp. 237–242
26 Goodman, S.N.: 'Toward evidence-based medical statistics. 1: The *P* value fallacy', *Ann. Intern. Med.*, 1999, **130**, pp. 995–1004
27 Stone, M., Brooks, R.J.: 'Continuum regression: cross-validated sequentially constructed prediction embracing ordinary least squares, partial least squares and principal components regression', *J. R. Stat. Soc. B, (Methodol.)*, 1990, **52**, pp. 237–269
28 Shi, Y., Eberhart, R.: 'A modified particle swarm optimizer'. The 1998 IEEE Int. Conf. on Evolutionary Computation Proc., 1998. IEEE World Congress on Computational Intelligence., 1998, pp. 69–73
29 Massey Jr. F.J.: 'The Kolmogorov–Smirnov test for goodness of fit', *J. Am. Stat. Assoc.*, 1951, **46**, pp. 68–78
30 Choqueuse, V., Yao, K., Collin, L., Burel, G.: 'Blind detection of the number of communication signals under spatially correlated noise by ICA and KS tests'. 2008 IEEE Int. Conf. on Acoustics, Speech and Signal Processing, ICASSP 2008, 2008, pp. 2397–2400
31 Wahlström, S.: 'Comparing forecasts from LSTAR and linear autoregressive models', 2004, available at http//:www.leg.ufpr.br
32 Granger, C.W., Terasvirta, T.: 'Modelling non-linear economic relationships' (OUP Catalogue, 1993)
33 Itoh, K.-I., Watanabe, S., Shih, J.-S., Sato, T.: 'Performance of handoff algorithm based on distance and RSSI measurements', *IEEE Trans. Veh. Technol.*, 2002, **51**, pp. 1460–1468

7 Appendix

In this Appendix, a numerical example is presented for illustration of PSO and OLS steps in LSTAR model building. In this example, a time series with ten samples and model order (p) of two and delay (d) of one is assumed. With these assumptions, (8) can be written as

$$y_t = \left(\phi_{10} + \phi_{11}y_{t-1} + \phi_{12}y_{t-2}\right)\left(1 - F_{\text{st}}(y_{t-1};\ \gamma,\ c)\right) + \left(\phi_{20} + \phi_{21}y_{t-1} + \phi_{22}y_{t-2}\right)F_{\text{st}}(y_{t-1};\ \gamma,\ c) + u_t \tag{35}$$

The SINR samples are set to

$$\begin{aligned} y_t &= \text{SINR(dB)} \\ &= [2.16,\ 13.51,\ 11.02,\ 7.48,\ 3.79, \\ &\quad 12.88,\ 12.39,\ 10.11,\ 7.33,\ 16.28] \end{aligned} \tag{36}$$

The delay of one means that modelling starts from the second SINR sample and the first sample is used in the calculation of

the parameters. Below, the steps of estimation of the parameters $\boldsymbol{\phi} = [\boldsymbol{\phi}_1^{\mathrm{T}}, \boldsymbol{\phi}_2^{\mathrm{T}}]^{\mathrm{T}}$ using PSO and OLS are explained.

1. The primary population with five particles (each one denoting the parameters c and γ) is generated randomly as

$$\boldsymbol{G} = \begin{bmatrix} \gamma & 2.19 & 1.53 & 1.46 & 1.22 & 4.68 \\ c & 10.78 & 5.46 & 5.31 & 2.23 & 6.41 \end{bmatrix}^{\mathrm{T}} \tag{37}$$

2. For each particle (each row of the matrix $\boldsymbol{G}$ including one pair of γ and c), we calculate $\boldsymbol{\phi} = [\boldsymbol{\phi}_{10}, \boldsymbol{\phi}_{11}, \boldsymbol{\phi}_{12}, \boldsymbol{\phi}_{20}, \boldsymbol{\phi}_{21}, \boldsymbol{\phi}_{22}]^{\mathrm{T}}$ using the OLS.

2.1. The calculations are explained for the first particle. For the first particle at the first iteration, we have $\gamma_1(1) = 2.19$ and $c_1(1) = 10.78$. $\tilde{\boldsymbol{w}}_t(1)$ is calculated for all times according to (16). For example, $\tilde{\boldsymbol{w}}_2(1)$ is calculated as follows

$$\boldsymbol{w}_2 = [1, \quad y_1, \quad 0]^{\mathrm{T}} = [1, \quad 2.16, \quad 0]^{\mathrm{T}} \tag{38}$$

$$\begin{aligned} F_{\mathrm{st}}(y_{t-1};\ \gamma_1(1),\ c_1(1)) &= \frac{1}{1 + \exp(-\gamma_1(1) \times (y_1 - c_1(1)))} \\ &= 6.34 \times 10^{-9} \end{aligned} \tag{39}$$

$$\begin{aligned} \tilde{\boldsymbol{w}}_2 &= \begin{bmatrix} w_2(1 - F_{\mathrm{st}}(y_{t-1};\ \gamma_1(1),\ c_1(1))) \\ w_2 F_{\mathrm{st}}(y_{t-1};\ \gamma_1(1),\ c_1(1)) \end{bmatrix} \\ &= \begin{bmatrix} \begin{bmatrix} 1 \\ 2.16 \\ 0 \end{bmatrix} \times (1 - 6.34 \times 10^{-9}) \\ \begin{bmatrix} 1 \\ 2.16 \\ 0 \end{bmatrix} \times 6.34 \times 10^{-9} \end{bmatrix} \\ &= [1 \quad 2.16 \quad 0 \quad 6.34 \times 10^{-9} \quad 1.37 \times 10^{-8} \quad 0]^{\mathrm{T}} \end{aligned} \tag{40}$$

Then, $\boldsymbol{\phi}$ is estimated by the OLS using (15) as

$$\begin{aligned} \hat{\boldsymbol{\phi}}_1(1) &= \left[\sum_{t=2}^{10} \tilde{w}_t(1)\tilde{w}_t^{\mathrm{T}}(1)\right]^{-1} \left[\sum_{t=2}^{10} \tilde{w}_t(1) y_t\right] \\ &= [15.68 \quad -0.64 \quad -0.12 \quad -2.29 \quad 1.07 \quad -0.08]^{\mathrm{T}} \end{aligned} \tag{41}$$

Table 13 MSEs of particles at the first iteration

Particle (k)	1	2	3	4	5
MSE	8.82	6.65	6.74	6.99	2.30

2.2. Fitness value (MSE in this paper) is computed using (18) as follows

$$\begin{aligned} \mathrm{MSE}_1(1) &= \frac{1}{n-d} \sum_{t=d+1}^{n} \varepsilon_t^2(k) = \frac{1}{9}\left(\varepsilon_2^2(1) + \cdots + \varepsilon_{10}^2(1)\right) \\ &= \frac{1}{9}\left(\left(y_2 - \tilde{w}_2^{\mathrm{T}}(1)\hat{\phi}_1(1)\right)^2 + \cdots + \left(y_{10} - \tilde{w}_{10}^{\mathrm{T}}(1)\hat{\phi}_1(1)\right)^2\right) = 8.82 \end{aligned} \tag{42}$$

2.3. If the fitness value of particle 1 is less than the best fitness value of that particle, that is, $pBest(1)$ in the history, the current values of particle 1 ($[\gamma_1(1), c_1(1)]$) are set as the new $pBest(1)$. In the first iteration, the history is empty and the current values are set as $pBest(1)$, but for the next iterations the best value is selected as above. Therefore $pBest$ for all five particles after the first iteration will be as (37).

3. The particle with the best fitness value (least MSE) among all particles is chosen as the $gBest$. The MSE values for all particles are listed in Table 13.

Therefore, the particle five is chosen as the $gBest$, that is, we have $gBest = [4.68\ 6.41]$.

4. The velocity and position of the particles are updated using (20) and (21), respectively, as

$$\boldsymbol{V}_2(k) = \begin{bmatrix} 2.37 & -2.46 \\ 2.69 & 0.64 \\ 2.72 & 1.34 \\ 2.72 & 3.73 \\ 0.24 & 0.10 \end{bmatrix}; \quad \boldsymbol{x}_2(k) = \begin{bmatrix} 4.56 & 8.32 \\ 4.22 & 6.10 \\ 4.18 & 6.65 \\ 3.94 & 5.96 \\ 4.92 & 6.51 \end{bmatrix} \tag{43}$$

If one of the maximum iteration number or minimum error criteria is attained, then the algorithm stops. In this example, the number of iterations and minimum error are assumed to be 10 and 0.1, respectively. The results for all iterations are listed in Table 14.

Finally, the LSTAR model of SINR data considering the parameters obtained at the tenth iteration is generated as

$$\begin{aligned} y_t = &\left(-2079.75 + 967.44y_{t-1} - 211.62y_{t-2}\right) \\ &\times \left(1 - F_{\mathrm{st}}(y_{t-1};\ 19.52,\ 7.06)\right) \\ &+ \left(-6.74 + 1.40y_{t-1} - 0.05y_{t-2}\right) \\ &\times F_{\mathrm{st}}(y_{t-1};\ 19.52,\ 7.06) + u_t \end{aligned} \tag{44}$$

doi: 10.1049/iet-spr.2014.0065

Table 14 Results of all iterations

Iterations	$gBset = [c\ \gamma]$	$\hat{\boldsymbol{\phi}} = \begin{bmatrix} \hat{\phi}_{10} & \hat{\phi}_{11} & \hat{\phi}_{12} \\ \hat{\phi}_{20} & \hat{\phi}_{21} & \hat{\phi}_{22} \end{bmatrix}$	MSE
1	[6.51 4.93]	$\begin{bmatrix} -764.44 & 359.54 & -78.70 \\ -25.69 & 2.78 & 0.20 \end{bmatrix}$	2.1326
2	[6.88 7.14]	$\begin{bmatrix} -315.48 & 152.05 & -33.33 \\ -20.22 & 2.38 & 0.13 \end{bmatrix}$	1.1493
3	[6.95 7.20]	$\begin{bmatrix} 9.00 & 2.21 & -0.64 \\ -28.54 & 2.98 & 0.24 \end{bmatrix}$	1.1599
4	[5.70 7.93]	$\begin{bmatrix} -5.23 \times 10^6 & 2.42 \times 10^6 & -5.28 \times 10^6 \\ -17.90 & 2.21 & 0.10 \end{bmatrix}$	0.9219
5	[7.21 10.91]	$\begin{bmatrix} -33.25 & 21.61 & -4.81 \\ -14.66 & 1.98 & 0.06 \end{bmatrix}$	0.6706
6	[7.18 12.63]	$\begin{bmatrix} -64.79 & 36.19 & -7.99 \\ -11.65 & 1.76 & 0.02 \end{bmatrix}$	0.5174
7	[7.15 14.36]	$\begin{bmatrix} -135.34 & 68.79 & -15.13 \\ -9.66 & 1.61 & -0.01 \end{bmatrix}$	0.4458
8	[7.12 16.08]	$\begin{bmatrix} -306.51 & 147.90 & -32.42 \\ -8.33 & 1.52 & -0.03 \end{bmatrix}$	0.4096
9	[7.09 17.80]	$\begin{bmatrix} -0.03 & 357.57 & -78.27 \\ -7.41 & 1.45 & -0.04 \end{bmatrix}$	0.3894
10	[7.06 19.52]	$\begin{bmatrix} -2079.75 & 967.44 & -211.62 \\ -6.74 & 1.40 & -0.05 \end{bmatrix}$	0.3773