Extended and revised version of VTC 1999-Fall paper

QUADRATIC AND LINEAR FILTERS FOR MOBILE RADIO CHANNEL PREDICTION

Torbjörn Ekman[1,*], *Gernot Kubin*[2], *Mikael Sternad*[1] *and Anders Ahlén*[1]

[1] Signals and Systems, Uppsala University, PO Box 528, SE-751 20 Uppsala, Sweden.
Email: te@signal.uu.se, ms@signal.uu.se, aa@signal.uu.se,
[2] Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology
Gusshausstrasse 25 /389, A-1040 Vienna, Austria. E-mail:g.kubin@ieee.org

ABSTRACT

Using a simple channel model for the multipath mobile radio channel, a nonlinear approach to long-range prediction of fading radio channels is motivated in a scenario with scatterers close to the mobile station.

The performance of linear and quadratic predictors is evaluated on wideband (6.4 MHz) measurements in the 1800 MHz band. The predictors are applied on estimated complex channel taps to find practical measures on the achievable prediction gain for different prediction horizons. Quadratic predictors, in the form of quadratic Volterra filters, are found to have good modeling properties, but very poor generalization properties as compared to linear predictors.

1. INTRODUCTION

Small-scale (fast) fading occurs when a mobile station travels through an interference pattern, generated by scattered and reflected propagating waves. Within some local area, the scattering geometry can be assumed to be time invariant. A geometry dependent low-dimensional deterministic mapping may thus exist from the channel transmission properties measured at one location to those at an adjacent location, when the mobile has moved a short distance. In this paper we motivate the use of linear and nonlinear long-range predictors using a simple physical model of the channel.

A common linear approach assumes that a large number of horizontal plane waves with vertical polarization arrive at the receiver from different directions in a plane [1]. Based on this assumption a spectral estimation approach [2] to long-range prediction can be taken. A linear predictor structure becomes less appropriate when the effect of non-planar waves are taken into consideration. Neural nets have been proposed for power prediction in wideband systems [3] and have been demonstrated to perform better than linear predictors [4] for 1 ms prediction horizons.

The ability to predict the channel is important for adaptive resource allocation and power control but could also be used for adaptive coding and modulation [5], [6]. For scheduling and adaptive resource allocation longer prediction horizons, on the order of 5-10 ms are of interest [7]. In this paper we therefore test linear and quadratic predictors on complex channel taps of measured channel impulse responses to find practical measures on the achievable prediction gain (PG) for different prediction intervals.

2. PHYSICAL MODELING OF THE FADING CHANNEL

The time-varying impulse response $c(\tau, t)$ for a baseband channel at time t in a multipath environment can be described by

$$c(\tau, t) = \sum_n a_n(t) e^{j(\varphi_n(t) - 2\pi f_c \tau_n(t))} \delta(\tau - \tau_n(t)), \quad (1)$$

where f_c is the carrier frequency and $a_n(t)$ is a time-varying attenuation factor covering antenna effects, path loss and attenuation due to reflection and scattering for the nth path [8]. The phase shift caused by reflectors and scatterers is described by $\varphi_n(t)$ and $\tau_n(t)$ denotes the propagation delay for the nth path. Let $g(\cdot)$ be a time invariant impulse response due to pulse shaping and receiver filtering and let the symbol interval be T. The discrete-time channel impulse response can then be described as an FIR-filter with the kth tap given by [9]

$$\begin{aligned} h_k(t) &= \int_0^{NT} g(Tk - \tau) c(\tau, t) d\tau \\ &= \sum_n g(Tk - \tau_n(t)) a_n(t) e^{j(\varphi_n(t) - 2\pi f_c \tau_n(t))}, \quad (2) \end{aligned}$$

where NT covers the length of the continuous-time impulse response. With an effective support of $g(\cdot)$ on the closed interval $[-MT, MT]$, the number of reflectors and scatterers

*This work was supported by the Swedish National Board for Industrial and Technical Development, the Swedish Research Council for Engineering Science and the Marie Curie Research Training Grant. We thank Ericsson Radio Systems AB for supplying measurement data.

contributing to the kth tap will be limited to paths with delays in the interval $[\max(0, T(k-M)), T(k+M)]$.

If we just take ray optics into account, omitting the effects of diffraction and Fresnel optics, a scatterer can be modeled as a secondary source induced by a wave front whereas a reflector generates a mirror image of the emitting source. Scatterers and the mirror images can be viewed as secondary sources. The path delay $\tau_n(t)$ can be decomposed into the sum of a time-varying delay from the secondary source to the mobile ($\tau_n^{MS}(t)$) and a time-invariant path delay from the base station to produce the secondary source, τ_n^{BS}. The scatterers and reflectors are assumed to be fixed. The phase shift due to the time invariant delay, τ_n^{BS}, can be included in a complex attenuation factor

$$\alpha_{n,k}(t) = g(Tk - \tau_n(t)) a_n(t) e^{j(\varphi_n(t) - 2\pi f_c \tau_n^{BS})}. \quad (3)$$

The discrete channel (2) can now be expressed as

$$h_k(t) = \sum_n \alpha_{n,k}(t) e^{j\phi_n(t)}, \quad (4)$$

where the remaining phase function is

$$\phi_n(t) = -2\pi f_c \tau_n^{MS}(t) = -\frac{2\pi}{\lambda} ||\mathbf{x}_n^{MS}(t)||. \quad (5)$$

Here $\mathbf{x}_n^{MS}(t)$ is a vector in space pointing from the nth secondary source to the mobile station (see Fig. 1) and λ is the wavelength. For spherical waves the position vector is parallel to the wave vector, $\mathbf{k}_n(t)$. The phase is solely a function of the electrical distance to the secondary source. When the distance changes by as little as one wavelength the phase $\phi_n(t)$ changes by 2π, causing the effect of small-scale (fast) fading. Consider the simplest mobile dynamics, that is, a

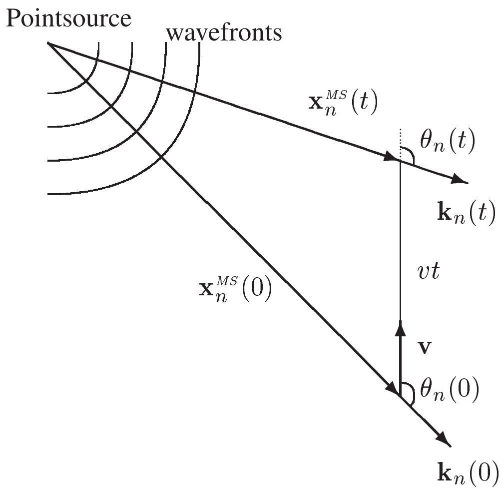

Figure 1: The change of path distance, $\mathbf{x}_n^{MS}(t)$, when moving with constant velocity, $\mathbf{v}$.

straight-line motion at constant velocity $\mathbf{v}$. For notational convenience we denote the initial position by $\mathbf{x}_n^{MS}(0) = \mathbf{x}_n^{MS}$ without explicit time index. We then have the phase function

$$\phi_n(t) = -\frac{2\pi}{\lambda} ||\mathbf{x}_n^{MS}(t)|| = -\frac{2\pi}{\lambda} ||\mathbf{x}_n^{MS} + \mathbf{v}t||. \quad (6)$$

Let $\theta_n(t)$ denote the angle between the position vector $\mathbf{x}_n^{MS}(t)$ (or wave vector $\mathbf{k}_n(t)$) and the velocity vector $\mathbf{v}$ and let $\theta_n = \theta_n(0)$, $v = ||\mathbf{v}||$, $x_n^{MS}(t) = ||\mathbf{x}_n^{MS}(t)||$. Then the phase function at the position $\mathbf{x}_n^{MS}(t)$ at time t can be rewritten as

$$\phi_n(t) = -\frac{2\pi x_n^{MS}}{\lambda} \sqrt{1 + 2\frac{vt}{x_n^{MS}} \cos\theta_n + \left(\frac{vt}{x_n^{MS}}\right)^2} \quad (7)$$

or, with $vt/x_n^{MS} \ll 1$ and a second-order approximation of $\sqrt{1+y} \approx 1 + y/2 - y^2/8$, we obtain

$$\phi_n(t) \approx -\frac{2\pi x_n^{MS}}{\lambda} \left[1 + \frac{vt}{x_n^{MS}} \cos\theta_n + \frac{1}{2}\left(\frac{vt}{x_n^{MS}}\right)^2 \sin^2\theta_n\right]. \quad (8)$$

Beyond the linear increase of the phase, due to the conventional Doppler shift, we obtain a quadratic term originating from the sphericity of the wave fronts.

The instantaneous frequency, the time derivative of the phase function, is given by

$$\dot{\phi}_n(t) = -\frac{2\pi}{\lambda}\frac{d}{dt} ||\mathbf{x}_n^{MS}(t)|| = -\frac{2\pi}{\lambda} v \cos\theta_n(t), \quad (9)$$

or using the approximate expression (8)

$$\dot{\phi}_n(t) \approx -\frac{2\pi}{\lambda} \left[v\cos\theta_n + \frac{t}{x_n^{MS}} (v \sin\theta_n)^2\right]. \quad (10)$$

The instantaneous frequency can thus be described as an instantaneous Doppler shift depending on the momentary angle between velocity and wave vector (9) or approximately as the Doppler shift at time $t = 0$ and a time dependent correction (10).

Assuming that the incoming waves are plane, that is $\theta_n(t) = \theta_n$, and a time invariant attenuation $\alpha_{n,k}$ in (4), we obtain the following commonly used approximation [1]

$$h_k(t) = \sum_n \alpha_{n,k} e^{-j2\pi f_{D_n} t}. \quad (11)$$

with the Doppler frequency $f_{D_n} = v\cos\theta_n/\lambda$. This model has been used as a basis expansion model for blind equalization [10] and also for long-range prediction of mobile radio channels [2]. The maximum deviation from the model (11) occurs for a *transversal vehicle movement* at $\theta_n = \pi/2$, when the velocity vector is orthogonal to the direction of the incident wave at time $t = 0$ and the nominal Doppler shift vanishes. This deviation is accumulated into a phase error which defines the maximum time interval over which the linear deterministic model can be used as an approximation in the present spherical-wave case. Defining this time interval, $T_{\pi/2}$, as the the interval when the phase approximation error is less than 90°, $|\,\Delta\phi_n(t)|_{T_{\pi/2}}\,| = \pi/2$, we obtain

$$T_{\pi/2} = \sqrt{\frac{\lambda x_n^{MS}}{2v^2}}. \tag{12}$$

For this time span, the above second-order approximation (8) can be well justified as we have

$$\left.\frac{vt}{x_n^{MS}}\right|_{t=T_{\pi/2}} = \sqrt{\frac{\lambda}{2x_n^{MS}}} \ll 1 \tag{13}$$

where the last inequality holds because even a close point scatterer will be many wavelengths away from the mobile transceiver.

The path loss for a scattered or reflected path is proportional to the path distance as [11]

$$a_n(t) \propto (x_n^{MS}(t) \otimes x_n^{BS}(t))^{-\alpha/2} \tag{14}$$

where $x_n^{MS}(t), x_n^{BS}(t)$ denotes the distance from the nth scatterer/reflector to the mobile station and the base station respectively and α is the power attenuation exponent. In (14) $x_n^{MS}(t) \otimes x_n^{BS}(t)$ denotes $x_n^{MS}(t) \cdot x_n^{BS}(t)$ for scattering and $x_n^{MS}(t) + x_n^{BS}(t)$ for specular reflection. The power attenuation exponent corresponding to path loss in free space is $\alpha = 4$. Even fairly close scatterers will give contributions with amplitudes on the order of a magnitude lower than the specular reflections. Thus it is appropriate to use the linear model (11) in situations where there are direct line of sight (LOS) or large contributions from specular reflections. In all other cases the effect of spherical waves introduces a significant deviation from (11). This can be seen from an example with a carrier frequency of 1800 MHz where the vehicle moves at $v = 90$ km/h past a close point scatterer at a distance $x_n^{MS} = 10$ m, resulting in $T_{\pi/2} = 36$ ms. Thus, the linear deterministic model is certainly inadequate for an estimation window of 150 ms, as we have in our measurements. This is also true for prediction horizons of 10 to 30 ms, when there are close scatterers. However, this problem does not exist for more distant (primary and reflected) sources, say, at $x_n^{MS} = 250$ m and beyond where we obtain $T_{\pi/2} = 178$ ms.

Similar effects with phase functions containing quadratic terms occur when the vehicle accelerates or makes a turn.

3. PREDICTORS

Under the assumption that the taps of a channel can be described as a weighted sum of complex sinusoids (11), a linear FIR-filter can give perfect predictions [12]. Let t denote the discrete time index for the channel samples. Assuming past impulse response taps to be exactly known, the L step ahead predictor, using p past samples, can be expressed as

$$\hat{h}_k(t+L|t) = \sum_{n=0}^{p-1} b_n h_k(t-\tau_n), \tag{15}$$

where the spacing of the delays τ_n is a user choice. The predictor coefficients, b_n, can be optimized directly by e.g. the least squares (LS) method or from iterating the one step ahead predictor. The predictor may be sub-sampled, thus τ_n may differ from n. This enables the predictor to use a long memory and still limit the number of coefficients. With a one step ahead predictor and no sub-sampling in the predictor the estimation of the FIR predictor can be interpreted as spectral estimation using an AR-model of the process.

If we assume that there are significant contributions to the channel from close point scatterers, the performance of the linear prediction approach will be inadequate, cf section 2. This is a motivation for investigating an extension of the linear predictor. Here we will use the quadratic Volterra filter [13].

$$\begin{aligned} \hat{h}_k(t+L|t) &= a + \sum_{n=0}^{p-1} b_n h_k(t-\tau_n) \\ &+ \sum_{n=0}^{p-1}\sum_{l=n}^{p-1} c_{nl} h_k(t-\tau_n) h_k(t-\tau_l) \end{aligned} \tag{16}$$

The most general quadratic filter for complex signals includes not only the quadratic terms but also all possible combinations of conjugate products. In the examples studied we have found no need for these extra terms.

For both the linear and the quadratic predictor we use the LS method to estimate the parameters for the predictors. To avoid numerical problems, the data are normalized so that linear and quadratic terms are of the same order of magnitude. Still the covariance matrix for the regressors may be ill-conditioned, resulting in low numerical accuracy and poor generalization capability for the predictor. To avoid this we regularize the solution [14] by adding small positive constants to the diagonal of the estimated covariance matrix (diagonal loading) prior to the LS estimation.

As a measure of prediction quality we use the prediction gain (PG), which is defined as the ratio between the variance of $h_k(t)$ and the power of the resulting prediction error, $e_k(t) = h_k(t) - \hat{h}_k(t|t-L)$. That is,

$$G(L) = 10\log_{10}\frac{E\{(h_k(t) - m_{h_k})^2\}}{E\{e_k(t)^2\}}, \tag{17}$$

measured in dB. Here $m_{h_k} = E\{h_k(t)\}$ denotes the average value of the time series.

4. EXPERIMENTS AND RESULTS

4.1. Measurements

Wideband radio channel measurements were performed at 1880 MHz, at distances of 200 to 2000 m from the base station antenna placed on the roof of a high building. The mobile antenna was placed on a car driving in a suburban area mostly at non-line of sight. The vehicle speed varied between 30 and 90 km/h. In total 25 measurement runs were recorded at different positions. The measurements consist of a repeatedly transmitted sequence of length 109 μs, resulting in 156 ms received signal at each measurement location. The baseband sampling rate of the receiver was 6.4 MHz.

The impulse response of the channel is modeled by a 120 tap FIR-filter, with parameters estimated by the LS method, for each repeated sequence. The time-span covered by the FIR-filter is 18.75 μs. For the identification, the signal from a back-to-back measurement is used as reference. This procedure results in 1430 consecutive complex impulse response estimates at each measurement location. The channel is thus sampled at 9.14 kHz, whereas the highest Doppler frequency is about 160 Hz.

A rough estimate of the SNR in the measured channel taps can be obtained from the power delay profile (PDP). The PDP is the time average of the squared amplitude of the taps. For taps not receiving any signal energy, the only contribution to the PDP is the noise, causing a noise floor as seen in Fig. 2. The amplitude of the PDP for tap k consists of signal power and noise power, where the noise power approximately is given by the noise floor level.

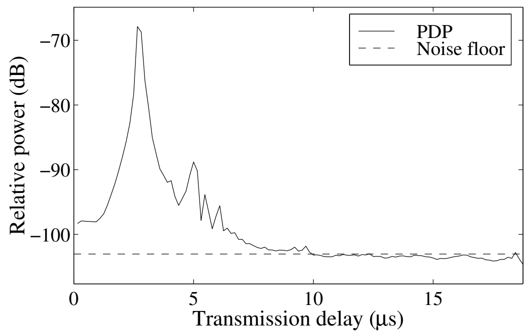

Figure 2: Power delay profile (PDP) and noise floor for one measurement location.

As the process is approximately band limited to $\pm v/\lambda$, and we have high oversampling, it is possible to reduce the noise level. A Butterworth low-pass filter of degree 4 with the cut-off frequency at the Doppler frequency for a vehicle speed of 110 km/h is a rough approximate model for the dynamics of the channel. Using the Butterworth filter together with the knowledge about the noise level, a smoother, with smoothing lag of 5 samples, is designed according to [15]. In an online application only the last available samples of the channel will need shorter smoothing lags.

4.2. Channel Simulator

A simple simulation model including the effects of scatterers in the vicinity of the mobile station can be derived from (4). The path losses are modeled as in (14) and the phase shifts φ_n in (3) are taken to be random variables uniformly distributed in $[0, 2\pi]$. The time-varying phases $\phi_n(t)$ are calculated as in (6).

The following simulation scenario is studied: A mobile station is passing three point scatterers which are located 5 m from the road and 5 m apart. The vehicle speed is 60 km/h, whereas the carrier frequency and channel sampling rate are selected as in the measurements described in subsection 4.1. No LOS or dominant reflector is assumed. Fig. 3 shows the instantaneous frequency for the different scattered wavefronts and the corresponding amplitudes.

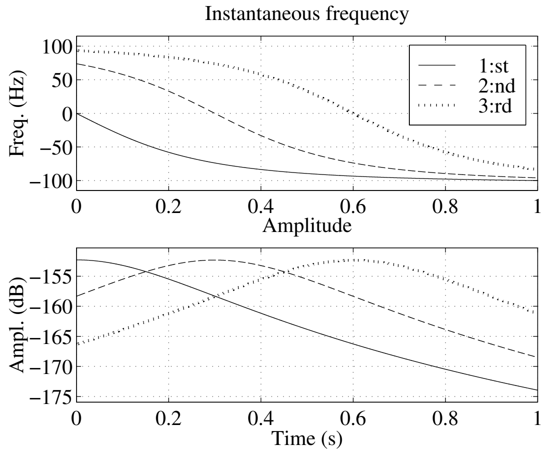

Figure 3: The instantaneous frequencies for 3 scatterers and their amplitudes. The maximum Doppler frequency is 104.4 Hz.

The performance of the noise reduction has been studied on the simulated data. Noise is added to give an SNR of 10 dB for the channel tap. After the smoothing filter the SNR has increased to 20 dB. The improvement decreases with higher SNR, e.g. for 30 dB the improvement is only 6 dB.

4.3. Prediction of Channel Parameters

For the simple simulation scenario with only three scatterers we evaluate the performance of a quadratic and linear predictor. Both use 9 delayed samples with delays evenly spaced over the interval $[0, 200]$ as seen in Fig. 4. This corresponds to a memory length of 22 ms and a sub-sampling

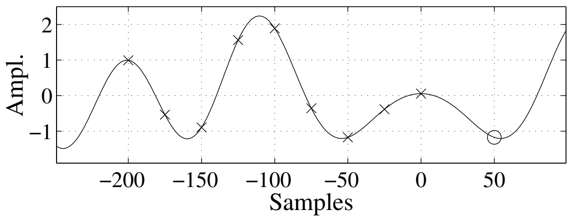

Figure 4: Short segment of a time series. 'x' denotes the samples used by the predictor and 'o' the predicted sample. The samples used for the predictor are spaced 25 samples apart, thus a sub-sampling factor of 25 is used.

factor of 25. The linear predictor and quadratic predictors use 9 and 55 complex parameters respectively. The prediction horizon is $L = 50$ samples (5.5 ms). The channel is normalized to have a variance equal to one and a very small noise term is added, resulting in an SNR of 120 dB. The first 2000 samples (220 ms) are used to build the models and the following 500 samples (55 ms) are used for validation. We use diagonal loading of the covariance matrix for the regressors with 10^{-5} for the quadratic terms and 10^{-9} for the linear. On the model set the PG for the linear and quadratic predictor is 22 dB and 36 dB respectively. For the validation set the prediction gains fall dramatically to 1 dB for the linear and to 6 dB for the quadratic predictor. The quadratic predictor shows high sensitivity to noise and to the choice of the diagonal loading term. When the noise level increases the quadratic predictor loses its generalization capability and fails in the validation.

A linear predictor using a sub-sampling factor of 5 and the same length of memory (22 ms), thus using 41 delayed samples and complex parameters, achieves higher PG both in the modeling set (38 dB) and in the validation set (12 dB). The large difference between the PG for the model and validation set for the linear predictors is due to the nonlinearity caused by the quadratic phase terms. Neither the linear nor the quadratic predictor seems to be able to model this nonlinearity. The results are summarized in table 1.

Pred.	Subsamp.	No. param.	PG mod.	PG val.
Quad.	25	55	36	6
Linear	25	9	22	1
Linear	5	41	38	12

Table 1: Table of the results for the different predictors on the simulated data. The prediction horizon is 50 samples.

We then investigated prediction of the channel coefficient-estimates from measured broadband data at 1880 MHz. For the measured impulse responses only those taps with SNR above 10 dB are used to evaluate the performance of the predictors. The noise level is reduced using a smoothing filter and, from the previous simulation results, we can assume that the effective SNR is at least 20 dB. The same predictors as above were used on the measurements, except that no diagonal loading was used for the linear terms. We vary the prediction horizon L from 25 to 100 samples, that is 2.7-11 ms. Of the 1430 samples available in the measured time series 10 are lost due to the smoothing (5 due to the smoothing lag and another 5 due to the initialization of the filter). Of the remaining 1420 samples the first $1200 + L$ are used for modeling and the remaining $220 - L$ samples constitute the validation set.

A total of 532 taps are used in the study. We average the achieved PG (in dB) for all the used taps and compare the PG for model and validation data in Fig. 5. We see that

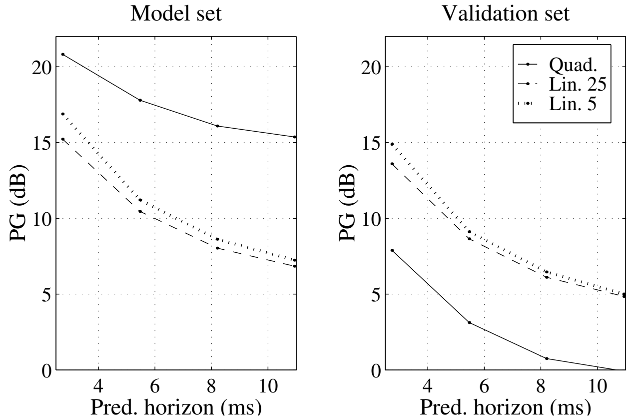

Figure 5: Average PG for different prediction horizons for both model and validation set with sub-sampling factor 25 and 5 respectively. The quadratic predictor has subsampling factor 25.

even though the quadratic predictor gives a much higher PG on the modeling set it fails to generalize to the validation set, whereas the linear predictors loose only 2 dB. There is no major difference in performance between using 41 or 9 parameters (sub-sampling factor 5 instead of 25) for the linear prediction. Neither does an increase of the memory to 300 samples (33 ms), using subsample factors 5 and 25, result in a significant improvement of the performance of the linear predictors on these data sets.

5. DISCUSSION

The LS solution for quadratic filters relies on accurate estimates of fourth order moments. This limits the use of quadratic filters in a time-varying environment such as a mobile radio channel. A single chirped signal

$$x(t) = \beta e^{i(\omega t+\zeta t^2)}, \tag{18}$$

can be modeled by the following deterministic recursion.

$$x(t) = e^{2\varsigma} \frac{x(t-1)^2}{x(t-2)}. \tag{19}$$

A quadratic Volterra model is thus unsuitable to model this behaviour.

The rapid fall off of predictability using the linear filters, losing 5 dB when going from prediction of 25 samples ahead (2.7 ms) to 50 samples (5.5 ms), indicates that the performance of block-adaptive linear predictors is limited for long-range prediction. However, the small loss in performance when going from the model set to the validation set indicates good generalization properties for the linear predictors. As the increase of the number of linear parameters hardly improves the performance, even the simple linear predictor exploits most of the available linear dependencies.

Possible extensions of the current work is to use nonlinear ARMAX models (NARMAX) and adaptive linear filters. Such models will be considered in our future work.

6. REFERENCES

[1] Jr W.C. Jakes, "Multipath interference," in *Microwave Mobile Communications*, Jr W.C. Jakes, Ed. IEEE, Piscataway, NY, 1974.

[2] A. Duel-Hallen T. Eycoz and H. Hallen, "Deterministic channel modelling and long range prediction of fast fading radio channels," *IEEE Communications Letters*, vol. 2, no. 9, pp. 254–256, Sept. 1998.

[3] X.M. Gao, X.Z. Gao, J.M.A. Tanskanen and S.J. Ovaska, "Power control for mobile DS/CDMA systems using a modified Elman neural network," in *Proc. IEEE* 47^{th} *Vehic. Tech. Conf.*, Phoenix, AZ, 1997, vol. 2, pp. 750–754.

[4] X.M. Gao, J.M.A. Tanskanen and S.J. Ovaska, "Comparison of linear and neural network-based power prediction schemes for mobile DS/CDMA systems," in *Proc. IEEE* 46^{th} *Vehic. Tech. Conf.*, Atlanta, GA, May 1996, pp. 61–65.

[5] S. Chua and A. Goldsmith, "Variable-rate variable-power MQAM for fading channels," in *Proc. IEEE* 46^{th} *Vehic. Tech. Conf.*, May 1996, pp. 815–819.

[6] S. Sampei T. Ue and N. Morinaga, "Adaptive modulation packet radio communication system using NP-CSMA/TDD scheme," in *Proc. IEEE* 46^{th} *Vehic. Tech. Conf.*, May 1996, pp. 416–420.

[7] N. C. Ericsson, "Adaptive modulation and schedueling of ip-trafic over fading channels," in *Proc. IEEE* 50^{th} *Vehic. Tech. Conf.*, Sep. 1999.

[8] John G. Proakis, *Digital Communications*, McGraw–Hill, New York, NY, second edition, 1989.

[9] M. Cedervall B.C. Ng and A. Paulraj, "A structured channel estimator for maximum-likelihood sequence estimation," *IEEE Communications Letters*, vol. 1, no. 2, pp. 52–55, March 1997.

[10] G.B. Giannakis and C. Tepedelenlioğlu, "Basis expansion of models and diversity techniques for blind identification and equalization of time-varying channels," *Proc. IEEE*, vol. 86, no. 10, pp. 1969–1998, Oct. 1998.

[11] A.F. Molisch J. Fuhl and E. Bonek, "Unified channel models for mobile radio systems with smart antennas," *IEE Proc.-Radar, Sonar Navig.*, vol. 145, no. 1, pp. 32–41, Feb. 1998.

[12] P. Stoica and R. Moses, *Introduction to Spectral Analysis*, Prentice Hall, NJ, 1997.

[13] G. L. Sicuranza, "Quadratic filters for signal processing," *Proc. IEEE*, vol. 80, no. 8, pp. 1263–1285, Aug. 1992.

[14] L. Ljung, *System Identification*, Prentice Hall, NJ, second edition, 1999.

[15] A. Ahlén and M. Sternad, "Wiener filter design using polynomial equations," *IEEE Trans. Signal Processing*, vol. 39, pp. 2387–2399, Nov. 1991.