Power Prediction on Broadband Channels

Mikael Sternad, Torbjörn Ekman, Anders Ahlén
Signals and Systems, Uppsala University, PO Box 528, SE-751 20 Uppsala, Sweden.
{ms,te,aa}@signal.uu.se

Abstract

The fading envelope in mobile communications causes rapid variations of the receiving conditions. Fast radio resource allocation and planning require accurate predictions of the changing received power. To form a prediction of the fading power, the individual taps of the channel are here predicted and their squared magnitude are summed. The dynamics of the taps can be modeled as subsampled autoregressive or ARMA-processes, leading to the different linear predictors studied in this paper. The prediction performance is evaluated on measured mobile radio channels at 45 locations with a bandwidth of 6.4 MHz and Doppler frequencies in the range 30-110 Hz. With the proposed predictors the normalized mean square error for the power prediction is below one percent for prediction ranges up to 2.5 ms, for a majority of the measured channels.

1. Introduction and outline

Prediction of rapidly time-varying mobile radio channels is of interest in e.g. power control, fast link adaptation [1], transmit diversity based on feedback, and fast resource allocation [2, 3, 4]. For example, power control loops in WCDMA include a feedback delay of 0.67 ms. This corresponds to a small fraction of the distance between two fading dips, and prediction over such horizons is relatively easy.

Fast radio resource allocation and planning would require more long-range predictions with adequate accuracy (within 2-3dB). The longer the horizons, the better resource optimization can be achieved.

Long-range prediction based on past observations is receiving increasing interest and some approaches are described in [5, 6, 7, 8]. The difficulty increases with the prediction length. Prediction over a large fraction of a Doppler wavelength is a very challenging problem. Model experiments offer some insights [7, 8, 9], but tests on measured data are essential.

We utilize channel sounding measurements acquired in suburban as well as urban areas at 1880 MHz, using a 6.4 MHz sampling rate[1]. FIR channel estimates of length 120 samples were obtained by least squares estimation over blocks of 700 samples.

We study long term power prediction over horizons of 1-10 ms, which correspond to up to one wavelength in space. While complexity is important, prediction performance will be the main focus here. Our previous studies on measured data [6, 7] have led to several conclusions that are essential for good performance:

1. The signal power changes abruptly around fading dips, while the channel taps vary smoothly. Power prediction based on power samples is therefore inferior to *predicting the individual taps* and then summing their squared magnitudes.

2. In highly oversampled channels, it is worthwhile to *filter the data to reduce noise* before prediction is performed. This is the case in our data set with a channel sampling frequency of 9.14 kHz and a highest Doppler frequency of 110 Hz, see Section 3.

3. Linear predictors have good generalization properties beyond the data set used for adjustment, while nonlinear channel prediction by Volterra filters or Multiple Adaptive Regression Splines fail in in this respect [6, 7]. We thus focus on prediction by *linear filtering of past channel tap estimates.*

4. A good prediction performance will depend on accurate modeling of the fading statistics. The fading environment will change due to movement of the mobile and this limits the length of data windows. Good accuracy is provided by *parametric time-series models* such as autoregressive (AR) or autoregressive-moving average (ARMA) models. These models should be *subsampled*, see Section 4, but still be trained on *all* samples within the adaptation window.

We here discuss two algorithms, described in Sections 4-5, that fulfill the above requirements:

Sub-sampled FIR predictors, that correspond to AR modeling of the Doppler spectra are introduced in Section 5.1. Sub-sampled Kalman prediction based on

[1] We thank Ericsson Research in Kista for supplying the measurement data.

ARMA modeling is discussed in Section 5.2.

A performance evaluation based on the measured data is summarized in Section 6.

Notation: Baseband samples (6.4 MHz) are indexed by n, sequences at the channel sampling rate (9.14 kHz) by t, while m-step sub-sampled sequences ($9.14/m$ kHz) are indexed by i. In filter polynomials

$$P(q^{-1}) = p_0 + p_1 q^{-1} + \ldots + p_{n_P} q^{-n_P} \quad , \qquad (1)$$

the backward shift operator ($q^{-\ell}x(t) = x(t-\ell)$) is used. Subscripts s denote subsampled signals and filter polynomials (1) designed for sub-sampled data.

2. Channel estimation

The data was recorded at 1880 MHz in central Stockholm as well as in Kista, a suburb of Stockholm. The mobile antenna was placed on a car driving at 20-60 km/h and the distances to the base stations varied between 200-2000 m, mostly without line of sight. The measurements resulted in 156.4 ms recordings of the baseband signal at 45 locations. The fading widths varied between 2 dB and 16 dB, with 7 dB on average.

For each location, time-varying parameters $h_k(n)$ in

$$r(n) = \sum_{k=1}^{n_k} h_k(n)u(n-k) + w(n) \qquad (2)$$

were estimated. Here, $r(n)$ is the received baseband signal sampled at 6.4 MHz, $h_k(n)$ represent complex-valued channel taps and $w(n)$ is noise.

Above, $u(n)$ is a known filtered transmitted sequence, obtained from back-to-back measurements of the receiver connected to the transmitter. The taps $h_k(n), h_i(n), k \neq i$ will thus be almost uncorrelated. A channel length $n_k = 120$, corresponding to a time span of 18.75 μs, was found to be sufficient at all measurement locations. The channel was assumed time-invariant over blocks of 700 samples (channel sampling period $T = 109.4$ μs ≈ 0.01 wavelength), an assumption that introduces negligible errors at these fading rates [10]. Block least squares estimates of the channel taps were calculated, resulting in 1430 channel samples

$$\bar{h}_k(t) = h_k(t) + v(t) \ ; \ k = 1, \ldots, 120 \ ; \ t = 1, \ldots, 1430 \ ,$$

covering 156 ms at each location. Here, $h_k(t)$ denote the noise-free tap samples, while $v(t)$ is zero mean noise that is uncorrelated with $h_k(t)$.

At each location, we select a subset ranging from one to 35 of the 120 taps that contribute at least 90% of the signal power. The power to be predicted is defined as the sum of squared magnitudes of these taps. The SNR ranges from 9 to 50 dB for the significant taps.

3. Wiener smoothing of noisy taps

The tap sequences $h_k(t)$ are approximately band-limited so the noise outside of their passbands can be reduced. Realizable IIR Wiener smoothers designed according to [11] provide good disturbance rejection, combined with low delay. They are based on ARMA models for the signal $h_k(t)$ and the noise $v(t)$.

As signal model for *all* taps, we use a Chebyshev type one low-pass filter of degree 4 with cut-off frequency corresponding to a vehicle speed of of 105 km/h (the model in Figure 1). The noise is well approximated as white, and its variance can be estimated from the power delay profile at each location.

Fixed-lag smoothers $\mathcal{F}(q^{-1})$ with lag $\ell = 5$ have been designed for each tap[2], using the estimated noise variance. The resulting noise-reduced time series is

$$y(t) = \mathcal{F}(q^{-1})(h_k(t) + v(t)) \approx \underbrace{h_k(t-\ell)}_{s(t)} + \underbrace{\mathcal{F}(q^{-1})v(t)}_{n(t)} \ , \qquad (3)$$

since the effect of the smoother on the noise-free tap is essentially a delay equal to the smoothing lag ℓ. The noise reduction results in an increase of roughly 10 dB in the SNR. It improves the attainable prediction error variance by 2 dB for the considered algorithms, as compared to operating on $\bar{h}_k(t)$ directly with predictor algorithms tuned to handle more noise [7].

Figure 1 illustrates the spectrum for one tap before and after smoothing.

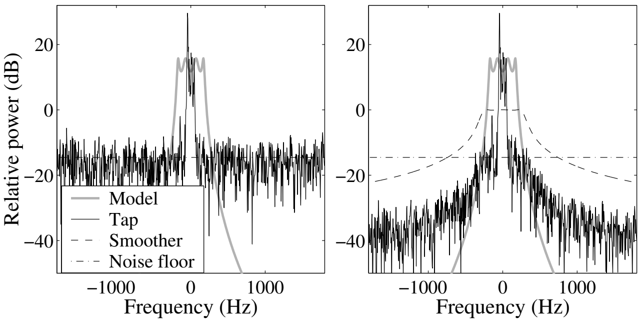

Figure 1: Power spectrum of the fixed fading model used for smoother design (grey) and the Doppler spectrum for a tap $\bar{h}_k(t)$ with SNR 14 dB. To the right, the tap $y(t)$ (solid) is shown after noise reduction with the smoother (dashed). The total bandwidth is ±4.6 kHz.

[2] A smoothing lag ℓ of 5 will for these data provide noise reduction only a few dB less than for unlimited lags.

The task is now to design predictors that use noise-reduced tap estimates to predict the noise-free taps

$$\hat{s}(t+L|t) = f(\{y(t-j)\},t) \quad , \; j \geq 0 \quad . \tag{4}$$

The effective prediction horizon will be $(L-5) \times 0.109$ ms, due to the smoothing delay $\ell = 5$ in (3). Since $s(t)$ is unknown, the prediction error is evaluated with respect to $y(t)$. The error can then of course not have variance below the variance of the noise $n(t)$.

4. Doppler spectrum estimation

The predictor (4) will be related to the Doppler spectrum of the tap $s(t)$, see Section 5. Predictors and spectral models should have sub-sampled form, i.e. $j = 0, m, 2m, \ldots$ with $m > 1$ in (4), for several reasons:

1. A longer filter memory can then be covered by a fixed number of parameters. This increases the modeling accuracy for low-frequency oscillations [8].

2. Spectral estimates obtained by prediction error methods normally minimize the variance of the one-step error [12]. With sub-sampling, we minimize the m-step error, which is more relevant to our L-step horizon.

3. The estimation of AR and ARMA models for narrow-band signals becomes better conditioned.

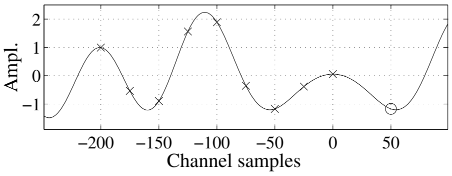

Figure 2: Sub-sampling with $m = 25$ (x). The ring indicates the sample to be predicted ($L = 2m = 50$).

Sub-sampling must be handled with some care, since it will introduce aliasing. It should be applied only to low-pass-filtered sequences. Furthermore, sub-sampled filters have periodic spectra, repeated m times, that can cause problems if their input has significant high-frequency content.

The smoother $\mathcal{F}(q^{-1})$ acts as our anti-aliasing low-pass filter which makes it safe to sub-sample the sequence $\{y(t)\}$ of noise-reduced taps. The noise-free tap $s(t)$ will be band-limited. It can safely be subsampled up to a rate of $m = 20$, giving a new Nyquist frequency of 4.6 kHz/20=230 Hz.

For $t = im, i = 1, 2, ..,$ the sub-sampled signal $y_s(i) = y(im)$ is

$$y_s(i) \approx s_s(i) + n_s(i) \quad ; \quad E|n_s(i)|^2 = \lambda_n \quad . \tag{5}$$

With $m > 7$, the noise $n_s(i)$ becomes approximately white, due to some remaining aliasing. An appropriate model structure for $s_s(i) = h_k(mi - \ell)$ is an ARMA model

$$s_s(i) = \frac{1 + c_1 q^{-1} + \ldots + c_{n_C} q^{-n_C}}{1 + d_1 q^{-1} + \ldots + d_{n_D} q^{-n_D}} e_s(i) \triangleq \frac{C_s(q^{-1})}{D_s(q^{-1})} e_s(i) \tag{6}$$

where $e_s(i)$ is zero mean white noise with variance λ_e.

The reason is as follows. A conventional model for a fading tap is based on N point reflectors and scatterers. It is expressed as a weighted sum of the outputs from N complex oscillators, which have slowly varying frequencies (Doppler shifts) [5]-[9].

We here instead suggest a linear time-invariant model that is valid for short intervals and is given by a sum of N *damped* oscillators:

$$\begin{aligned} x_{s,n}(i) &= \rho_n e^{j\omega_{D_n} T m} x_{s,n}(i-1) + w_{s,n}(i) \\ s_s(i) &= \sum_{n=1}^{N} \alpha_n x_{s,n}(i) + \epsilon_s(i) \quad . \end{aligned} \tag{7}$$

Here, Tm is the sub-sampling interval in seconds, ω_{D_n} is an average Doppler angular frequency for ray n and $\rho_n \leq 1$ introduces damping. This damping is due to averaging over slowly time-varying frequencies, as well as the presence of continuous scatterers. Uncertainties in the modeling are represented by $w_{s,n}(i)$ and $\epsilon_s(i)$, which are all band-limited, since $s_s(i)$ is band-limited, but are otherwise unknown. We can treat them as noise and substitute the resulting sum of stochastic processes in (7) by *one* system, with the same second order moments. This operation is called spectral factorization, and it results in an innovation model of the form (6), with $n_D = N$ and $n_C = N - 1$.

A similar ARMA model can also be used for $y_s(i)$

$$y_s(i) = \frac{\beta_s(q^{-1})}{D_s(q^{-1})} \varepsilon_s(i) \quad , \tag{8}$$

where $\varepsilon_s(i)$ is zero mean white noise with variance λ_ε. If $n_s(i)$ is assumed white, the spectra of $y_s(i), s_s(i)$ and $n_s(i)$ will be related via

$$\lambda_\varepsilon \frac{\beta_s(e^{-j\omega})\beta_s^*(e^{j\omega})}{D_s(e^{-j\omega})D_s^*(e^{j\omega})} \approx \lambda_e \frac{C_s(e^{-j\omega})C_s^*(e^{j\omega})}{D_s(e^{-j\omega})D_s^*(e^{j\omega})} + \lambda_n \tag{9}$$

A complication for the prediction problem is that only $y_s(i) = s_s(i) + n_s(i)$ is measurable rather than $s_s(i)$. If the noise level is low, the model (6) can, however, be approximated by the ARMA model (8) for $y_s(i)$.

Parameter estimation is performed for a limited data window, due to the (slow) time variability. The best accuracy is then obtained by prediction error methods applied to low order models with few zeros [12].

All channel samples $y(t)$ should be used in the estimation of (8). We create and utilize m sub-sampled data vectors by shifting the data in $y(t)$ by one step.

5. Prediction algorithms

5.1. FIR prediction based on AR modeling

It is well known that accurate estimation of zeros of ARMA models requires long data records [12]. A first attempt to predict $s(t)$ would therefore be to approximate (8) by an autoregressive model, $(\beta_s(q^{-1}) = 1)$. The least squares method is then used to estimate $D_s(q^{-1})$. A sub-sampled FIR predictor

$$\hat{s}_s(i+k|i) = \hat{y}_s(i+k|i) \overset{\Delta}{=} G_s(q^{-1})y_s(i) \qquad (10)$$

or

$$\hat{s}(t+L|t) = \hat{y}(t+L|t) = G_s(q^{-m})y(t) \quad ,$$

predicts $L = mk$ channel samples ahead. It can be optimized in an MSE sense by solving the identity

$$1 = D_s(q^{-1})F_s(q^{-1}) + q^{-k}G_s(q^{-1}) \qquad (11)$$

with respect to $F_s(q^{-1})$, of degree $k-1$, and $G_s(q^{-1})$ of degree n_D-1 [12]. This is called the *Direct FIR Method* in Section 6. Alternatively, the predictor coefficients can be obtained directly by using least squares i.e. by minimizing $E|y_s(i+k) - \hat{y}_s(i+k|i)|^2$.

Although these approaches are structurally incorrect $(\beta_s(q^{-1}) = 1, \hat{s}(t+L|t) = \hat{y}(t+L|t))$, they provide robust and rather accurate predictions, see Section 6.

5.2. Indirect ARMA-based predictor design

In Section 4, we argued that the best time-invariant model for a baseband channel tap is the (sub-sampled) ARMA model (6). We here present an indirect approach based on ARMA modeling and subsequent Kalman estimation, which proceeds as follows. First we estimate $\beta_s(q^{-1})$, $D_s(q^{-1})$, and λ_ε in (8) with a prediction error identification algorithm.

Stable β_s-polynomials, are required by Wiener and Kalman predictors. We therefore add a small term $\lambda_p > 0$ to the central coefficient of $\lambda_\varepsilon\beta_s\beta_s^*$. (This corresponds to adding a weakly colored noise to $n(i)$.) The left-hand numerator of (9) is thus substituted by

$$\lambda_g\gamma_s(e^{-j\omega})\gamma_s^*(e^{j\omega}) \overset{\Delta}{=} \lambda_\varepsilon\beta_s(e^{-j\omega})\beta_s^*(e^{j\omega}) + \lambda_p \quad .$$

Since D_s is known from the identification and λ_n can be determined from the smoothed channel sounding measurements, $\lambda_e C_s C_s^*$ could ideally be calculated by the use of spectral subtraction. However, unless we have a very accurate estimate of β_s, this may result in erroneous zeros of $C_s(z^{-1})$ located very close to the stability border. This in turn would deteriorate the prediction performance. As a crude estimate, we instead set $C_s(q^{-1}) \approx \gamma_s(q^{-1})$. An appropriate λ_e is then adjusted so that the variance of $y_s(i)$ becomes correct.

Based on estimates of $\gamma_s(q^{-1})$, $D_s(q^{-1})$, λ_e and λ_n a state -space model, such as (7), with (5) as the measurement equation, can be constructed. A set of m Kalman predictor is then designed to obtain $\hat{s}(t+L|t)$.

A similar procedure can be based on an AR model of $s_s(i)$. This is called the *Indirect AR Method.* It uses (8) with $\beta_s(q^{-1}) = 1$ as an approximation for (6).

6. Results

The prediction performance of three methods are now compared on measured complex taps. The direct approach uses a sub-sampled FIR-filter (10) derived via (11) from an LS-estimated AR-model for $y_s(i)$. The two indirect approaches use sub-sampled Kalman predictors, one based on an estimated AR-model for $s_s(i)$, the other on an estimated ARMA-model.

The variance of the sub-sampled measurement noise, λ_n, is used as a tuning variable for the Kalman predictors. It is set to the highest level in the sub-sampled smoothed noise spectrum.

After removing an initial transient, due to the smoothing operation, 1420 samples for each tap can be used. The first 1000 samples are used for identification of the AR- and ARMA-models. The following 420 channel samples are then used for evaluating the prediction performance. None of the channels showed significant time-variability within the training intervals.

A sub-sampling factor of $m = 8$ is chosen to illustrate the good short range power prediction properties. As the training interval is unfortunately short relative to the tap fading pattern, we can only use models of low order, with AR- and MA-polynomials of order $n_D = 4$ and $n_\beta = 2$ respectively.

6.1. Tap prediction performance

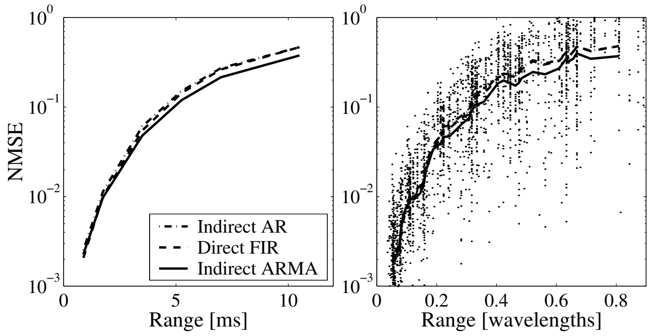

Figure 3: Medians of the normalized prediction MSE for all 407 significant taps at the 45 locations.

In Figure 3, the median of the prediction NMSE as a function of prediction range, expressed both in time and in traveled distance, is presented. The direct and indirect predictors show similar prediction performance, with the ARMA-based predictor giving the best performance. In all the figures the scatter plots show the performance of the ARMA-based predictor.

6.2. Received power prediction performance

We now calculate the sum of predicted powers of the significant taps. The medians of this power prediction NMSE's for all channels are presented in Figure 4. Two simple predictors are also included for comparison: The average power and the last available sample of the power. Their performances are much inferior for predictions up to 0.4 wavelengths, with few exceptions.

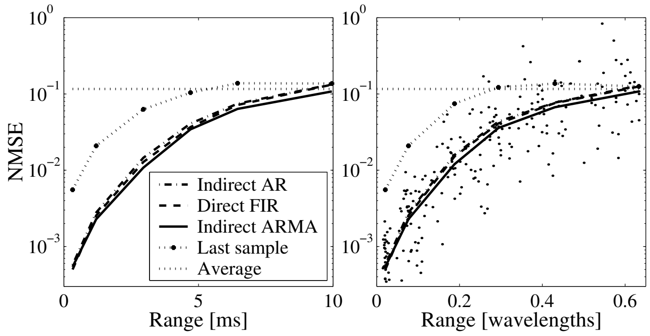

Figure 4: Medians of the normalized prediction MSE for the received power at 45 locations.

A more relevant measure of performance is how large portion of time the predicted power is within a certain bound relative to the true power. This is shown in Figure 5 for a 1 dB bound. For 1-7 ms predictions requiring a given percentage level, the indirect ARMA-predictor can on these channels be used for ranges more than 2 ms longer than the last sample predictor.

The obtained results are encouraging, but the performance presented here is essentially limited by the short data records, which limit the AR and ARMA estimation accuracy.

We believe that still better performance can be obtained with training windows longer than 0.1 ms. The length of the usable training window will ultimately be limited by the time-variability of Doppler spectra due to the movement relative to the closest scatterers.

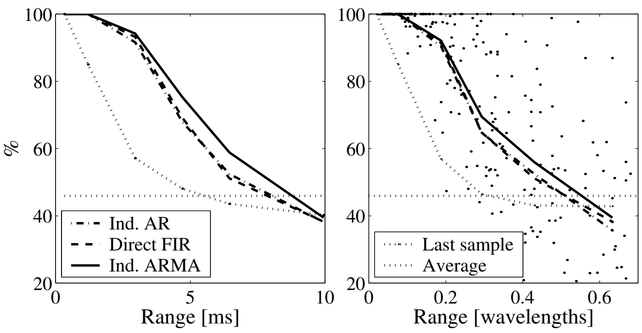

Figure 5: Percentage of time that the power prediction is within 1 dB of the measured power. Lines represent local medians over 45 points for different methods.

References

[1] S. Chua and A. Goldsmith, "Variable-rate variable-power MQAM for fading channels," *IEEE Vehich. Tech. Conf.*, Atlanta, GA, May 1996, pp. 815-819.

[2] N. C. Ericsson, "Adaptive modulation and scheduling of IP-traffic over fading channels," *IEEE Vehicular Tech. Conf. VTC'99-Fall*, Amsterdam, Sept. 1999, pp. 849-853.

[3] N. C. Ericsson, S. Falahati, A. Ahlén and A. Svensson, "Hybrid type-II ARQ/AMS supported by channel predictive scheduling in a multi-user scenario," *IEEE Vehicular Tech. Conf. (VTC2000-Fall)* , Boston, MA, USA, Sept. 24-28, 2000, pp. 1804-1811.

[4] M. Kawagishi, S. Sampei and N. Morinaga, "A novel reservation TDMA based multiple access scheme using adaptive modulation for multimedia wireless," *IEEE Vehicular Tech. Conf. VTC98* , 1998, pp. 112-116.

[5] J.B. Andersen, J. Jensen, S.H. Jensen and F. Fredriksen, "Prediction of future fading based on past measurements," *IEEE Vehich. Tech. Conf.- VTC'99-Fall*, Amsterdam, Sept. 1999, pp. 151-155.

[6] T. Ekman, G. Kubin, M. Sternad and A. Ahlén, "Quadratic and linear filters for mobile radio channel prediction," *IEEE Vehich. Tech. Conf.- VTC'99-Fall*, Amsterdam, Sept. 1999, pp. 146-150.

[7] T. Ekman, *Prediction of Mobile Radio Channels.* Licentiate Thesis, Signals and Systems, Uppsala University, Sweden, Dec. 2000.

[8] A. Duel-Hallen, S. Hu and H. Hallen, "Long-range prediction of fading signals," *IEEE Signal Processing Magazine* vol. 17, pp. 62-75, May 2000.

[9] R. Vaughan, P. Teal and R. Raich, "Short-term mobile channel prediction using discrete scatterer propagation model and subspace signal processing algorithms," *IEEE Vehich. Tech. Conf.- VTC'00-Fall*, Sept. 2000.

[10] T. Ekman, "Analysis of the LS estimation error on a Rayleigh fading channel," *IEEE Vehich. Tech. Conf.- VTC'01-Spring*, Rhodes, Greece, May 2001.

[11] A. Ahlén and M. Sternad, "Wiener filter design using polynomial equations," *IEEE Transactions on Signal Processing*, vol. 39, pp. 2387-2399, 1991.

[12] L. Ljung, *System Identification: Theory for the User.* Prentice-Hall, Second ed. 1999.