

# A new look at the body area network channel model

David B. Smith<sup>†</sup>, Leif W. Hanlen and Tharaka A. Lamahewa  
NICTA<sup>‡</sup> and the Australian National University,  
<sup>†</sup>Locked Bag 8001, Canberra, ACT 2601, Australia  
David.Smith@nicta.com.au

**Abstract**—A new approach to compare models for body area networks (BAN) that accommodates multiple links and multiple subjects is presented. The absolute measure described allows comparison across a spectrum of possible characterizations; ranging from single-parameter for an entire ensemble, through to per-activity, per-subject and per-link based parameterized models. The use of an explicit trade-off between error and complexity, combined in a goodness-of-fit measure, are shown to have important consequences when applied to a range of typical BAN channel data. It is shown that there are different implications in choice of model, and it's associated complexity, for mixed-activity “everyday” data, when compared with set-activity dynamic data (e.g. walking). The deficiency of mean path loss, or even median path loss measures, as a sole characterization of the BAN channel is also highlighted.

## I. INTRODUCTION

Body Area Networks, with networks of sensor/actuators placed around the human body, represent the next generation of personal area networks. Understanding body-area propagation, and hence developing a model for the BAN channel, is key to future BAN implementations [1]–[8]. Development of an accurate BAN channel model is particularly motivated by the need for reliable, low power operation for BAN [9].

An ideal BAN channel model has a simple statistic that is accurate on any given data set, whilst sufficiently general to accommodate “everyday” variation. Models parameterized by a given activity (eg. walking, running, sitting) or given sensor placement (per-link), have some value for BAN design, particularly when seeking to characterize channel dynamics [6], [10], [11]. However model efficacy depends on apriori knowledge [12]. Further models applying one statistic to an agglomerate, e.g. combining particular activity measurements and particular per-link measurements, may lose understanding due to modeling error [6]. An “everyday” model is a compromise between these demands.

In order to characterize given dynamic activity data, with particular sensor placement; e.g. a chest to right-hip link with subject running, may requires many thousands of samples over periods of 10s or more [10], with high sampling rates of 1 kHz or more. In contrast to characterize “everyday” behavior, BAN measurements must have moderately high sampling rates (above 100Hz), and very long durations (periods of an hour or more). Further better characterization of everyday activity is facilitated by many simultaneous links and multiple

subjects. Thus data-sets may have millions of samples, with the complication that the agglomeration of such data sets is large. In consideration of the various BAN channel modeling requirements, in this work we make the following contributions:

- 1) A systematic channel modeling method is presented, that can be used to evaluate potential models of the BAN channel by a novel absolute metric. This goes beyond standard BAN channel characterization that is based on traditional statistical fitting and comparison.
- 2) The modeling tool described can be used to evaluate potential characterizations of varying complexity of set activity data, as well as being used to evaluate potential characterizations of “everyday” data; for a range of sensor placements.
- 3) The utility of a per-link reference histogram; as an error-free reference for complete data-sets, and as an upper-bound on complexity, is described and demonstrated. Model options of range of complexities are evaluated and compared. These provide a benchmark for future BAN channel model evaluation.

## II. CHANNEL MODEL CHARACTERIZATION

In order to choose the best characterization of data we propose the following goodness of fit function. This function, which represents a generalization of various criteria for model selection<sup>1</sup>, for a model with parameters  $\theta = \{\theta_1, \dots, \theta_p\}$  applied to data  $\mathbf{x}$  with  $n$  samples is:

$$\mathcal{G}\{\theta, \mathbf{x}\} \triangleq \mathcal{E}\{\theta, \mathbf{x}\} + \mathcal{C}\{\theta, \mathbf{x}\} \quad (1)$$

where  $\mathcal{E}\{\cdot\}$  is an increasing function of error between model and data, and  $\mathcal{C}\{\cdot\}$  is an increasing function of number of parameters. The “optimal” model minimizes  $\mathcal{G}\{\theta, \mathbf{x}\}$  which reduces both over-fitting of data and model error. The second order Akaike information criterion (AIC) [13] may be formulated under (1) as

$$\mathcal{G}_{\text{AIC}}\{\theta, \mathbf{x}\} = \underbrace{[-2 \ln(L(\hat{\theta}|\mathbf{x}))]}_{\mathcal{E}_{\text{AIC}}\{\theta, \mathbf{x}\}} + \underbrace{\left[2p + \frac{2p(p+1)}{(n-p-1)}\right]}_{\mathcal{C}_{\text{AIC}}\{\theta, \mathbf{x}\}} \quad (2)$$

where  $\ln(L(\hat{\theta}|\mathbf{x}))$  is the maximised log-likelihood, based upon the maximum-likelihood (ML) estimate of model parameters  $\theta$ , given the data  $\mathbf{x}$ .

<sup>1</sup>Hence this function is not limited to propagation data

<sup>‡</sup>NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy and the Australian Research Council through the ICT Centre of Excellence program.

Another common model selection criterion; the Bayesian Information Criterion, BIC, (or Schwarz criterion) [14], which is equivalent to the MDL criterion [15] for a large sample size [16], may be formulated under (1) as

$$\mathcal{G}_{\text{BIC}}\{\boldsymbol{\theta}, \mathbf{x}\} = \underbrace{[-2 \ln(L(\hat{\boldsymbol{\theta}}|\mathbf{x}))]}_{\mathcal{E}_{\text{BIC}}\{\boldsymbol{\theta}, \mathbf{x}\}} + \underbrace{[p \ln n]}_{\mathcal{C}_{\text{BIC}}\{\boldsymbol{\theta}, \mathbf{x}\}} \quad (3)$$

For BAN's with multi-link measurements, both the AIC approach and the BIC approach (and thus also the MDL approach) suffers from the following problems:

- 1) *Relative*: the AIC and the BIC only provides an ordering of models. For different data sets with different parameterizations it is meaningless to compare AIC (or BIC) values.
- 2) *Log-likelihood*: tends to ignore uncommon values – the models may over-estimate rare event probabilities [11].

#### A. Model Selection

The demonstrated form of both the BIC and the AIC can be utilised to develop an absolute measure by choosing a common reference against which other models are compared. A natural reference point is the joint empirical histograms of the data sets. That is, given  $M$  data sets, we choose  $B$  histogram bins, and for each set  $m = 1, \dots, M$  find the histogram  $H_m(b)$  with  $b = 1, \dots, B$ . This ‘model’ has  $M \times B$  free parameters. Assuming (very) large  $n$ , criteria for a goodness of fit include:

- 1) The comparison of any model against a reference histogram is a metric – the distance between the model pdf and the reference histogram, evaluated at the histogram bin centres
- 2) The number of parameters is given by the number of model options  $M$  (per set models), and the number of free-parameters per option  $p_m$ .

We will consider  $M$  univariate empirical histograms  $H_m$ ,  $m = 1 \dots, M$ , defined at  $B$  bin centres  $\beta_b \in \mathbf{x}$ ,  $b = 1, \dots, B$ , and  $M$  univariate (continuous) probability density functions  $0 \leq F_m(\mathbf{x}) < \infty$ , to define an absolute goodness of fit as

$$\mathcal{G} \triangleq \underbrace{\frac{1}{MB} \sum_{m,b} |H_m(\beta_b) - F_m(\beta_b)|^2}_{\mathcal{E}} + \underbrace{\log_2 \left( \sum_{m=1}^M p_m \right)}_{\mathcal{C}} \quad (4)$$

and the base-2 logarithm follows complexity suggestions of [17]. Eqn (4) satisfies the criteria of (1). We neglected the terms relating to  $n$ , as seen in the BIC and AIC, since  $n$  is uniformly large for best choices of channel model.

### III. DATA SETS ANALYSED

We use two body-area network (BAN) data sets for illustration:

- The “everyday” data set in [18] with  $M = 4$  links (left/right-hip  $\rightarrow$  right wrist; left/right hip  $\rightarrow$  right-ankle), a narrowband measurement at engaged in everyday activities of an adult male subject over a period of 9 hours. The data per-link contains over 2.9 million samples. Samples

of channel gain were recorded at 200 Hz using small wearable channel sounders/radios (enabling free movement). The radios included a small ceramic multilayer chip antenna from Phycomp which is omnidirectional in the H-plane, had a gain of 1.2 dBi, and was operational at 2.4 GHz.

- A “dynamic” data set described in [19], with  $M = 9$  links (right-wrist, right-ankle, back  $\rightarrow$  chest; left/right-wrist, left/right-ankle, back, chest  $\rightarrow$  right-hip). One set activity is for an adult male subject running on-the-spot, and an other set activity for walking on the spot. Each individual per-link, per-activity subset contains 20,000 samples over a period of 20 s, corresponding to a channel gain sampling of 1 kHz. (Thus 180,000 samples for walking, and also 180,000 samples for running). The measurements were made using a vector signal analyser, and the antennas used were Octane BW-2400-900 wearable flexible antennas. These antennas are approximately omnidirectional, with similar radiation patterns in both E- and H-planes.

### IV. RESULTS

We find empirical histograms (using  $B = 50$  bins) and statistical fits for individual data sub-sets (per-link) and ensembles (agglomerate) for both data-sets; for the “everyday” data set and for the “dynamic” data set. For the “dynamic” data set we separate agglomerates into walking and running activity (i.e. we don’t combine walking and running into one agglomerate). For distribution fitting, we chose the second-order AIC measure in (2), to choose between six standard statistical models for channel gain distributions; i.e. Normal, Lognormal, Gamma, Nakagami-m, Weibull and Rayleigh distributions. We note that all apart from the Rayleigh distribution have 2 parameters in their statistic, thus since in no cases was a Rayleigh fit found by (2), all distribution statistics found had 2 parameters.

#### A. “Everyday” data results

The recorded power profiles in terms of received signal strength (in dBm)<sup>2</sup>, for the “everyday” data set analysed here is shown in Fig. 1. Over a period of 9 hours, for which the sensors were placed on the subject, the range of everyday activities are categorised.

Fig. 2 shows error  $\mathcal{E}$  vs complexity  $\mathcal{C}$  for various model options for the everyday data. Equivalent goodness  $\mathcal{G}$  is given by  $\mathcal{E} + \mathcal{C} = \text{const}$  and ‘better’ models will appear closer to the origin. For the per-link histogram in Fig. 2, the number of parameters  $P = M \times B = 200$ . In all cases the lognormal distribution was the best-fit. The number of parameters for the *mean per-link & agglomerate stat* is  $P = M + 2$ , since there are  $M$  means, and 2 free parameters. Due to the inaccuracy of mean-per-link (with  $P = M = 4$  parameters), and given single mean and median, we expand lower portion of the graph to better show the relative error of other model options.

<sup>2</sup>Note that the transmit power for the “Every-data” data is 0 dBm

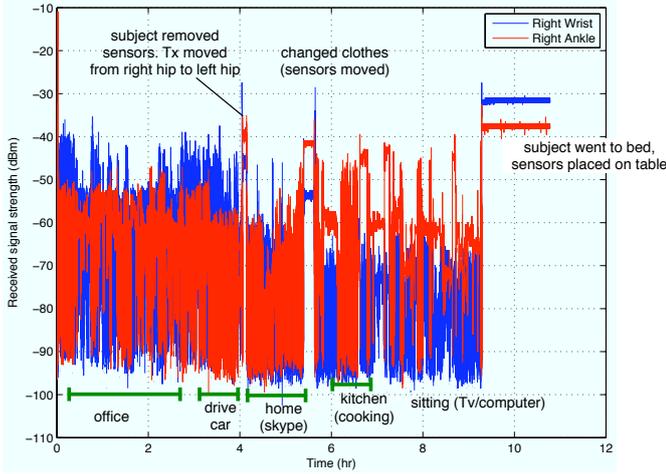


Fig. 1. Everyday measurement power profile and description

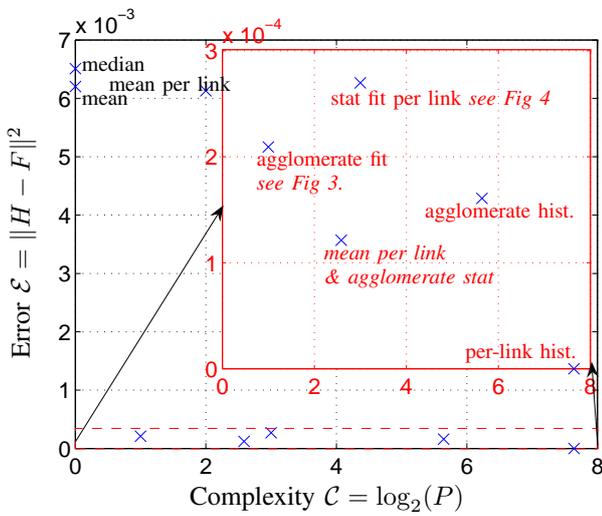


Fig. 2. Comparison of error  $\mathcal{E}$  and complexity  $\mathcal{C}$  for the everyday data. Note inset is the zoomed-in lower portion of the graph

The empirical histogram for all data sets gives zero error but excessive complexity  $P = MB$  (as expected). A combined histogram is also complex  $P = B$  and has moderate error. The error caused by using a simple mean as the model (which assumes average path-loss defines the model) is shown: the complexity is negligible but the error is large. In terms of goodness  $\mathcal{G}$  and as a trade-off between error  $\mathcal{E}$  and complexity  $\mathcal{C}$ , Fig. 2 shows that one of either (a) a mean-per-link with a lognormal statistic fitted to agglomerate data with mean-removed from each link; or (b) a lognormal fit to agglomerate data; is the preferable model. Option (a) is preferable in terms of  $\mathcal{E}$ , and (b) is preferable in terms of  $\mathcal{C}$ .

The statistical lognormal fits, for each link, and agglomerate data, with and without mean-per-link removed, are given in Table I. The lognormal fit to agglomerate data is shown in Fig. 3, where the fit is not perfect since the means of each data set are different, and the number of data sets is small. An example per-link fit for Left-Hip to Right-Wrist is shown in

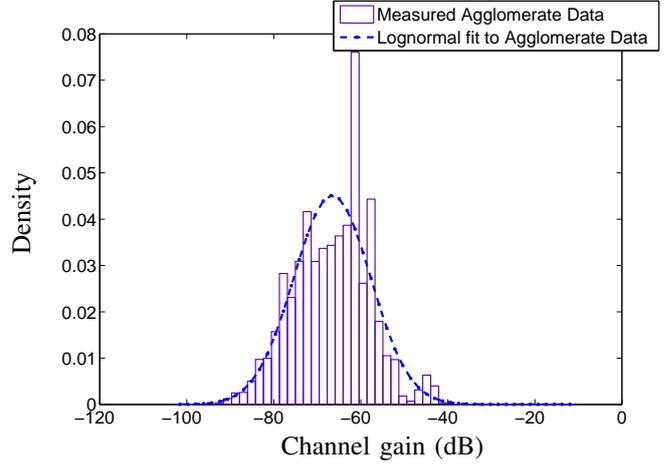


Fig. 3. Probability density, (empirical histogram) with best (Akaike) matched continuous distribution for everyday data.

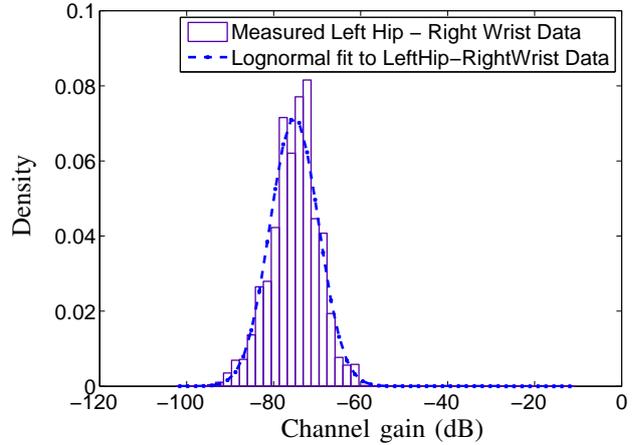


Fig. 4. Probability density, (empirical histogram) with best (Akaike) matched continuous distribution for a single-link everyday data.

Fig. 4.

### B. “Dynamic” set-activity data results

Figs. 5 and 6 shows error  $\mathcal{E}$  vs complexity  $\mathcal{C}$  for the same model options as for Fig. 2; for “Walking” in Fig. 5, and “Running” in Fig. 6, for the “Dynamic” set-activity data. For the per-link histograms for Figs. 5 and 6, the number of parameters is  $P = M \times B = 9 \times 50 = 450$ . The number of parameters for the *mean per-link & agglomerate stat* is  $P = M + 2 = 11$ , since there are  $M$  means, and 2 free parameters in both statistics for Fig. 5 and 6. Due to the

TABLE I  
BEST LOGNORMAL (LN) FITS TO EVERYDAY DATA; PARAMETERS:  $\mu$  IS LOG-MEAN,  $\sigma$  IS LOG-STANDARD DEVIATION. TRANSCIEVER (TX-RX) POSITIONS LEFT HIP-LH, RIGHT HIP-RH, RIGHT WRIST-RW, RIGHT ANKLE-RA. *Agg* - AGGLOMERATE DATA; *-mpl* - MEAN-PER-LINK REMOVED

Tx-Rx	LN parameters	Tx-Rx	LN parameters
RH-RW	$\mu = -7.07, \sigma = 0.77$	LH-RA	$\mu = -7.51, \sigma = 1.00$
RH-RA	$\mu = -7.25, \sigma = 0.73$	<i>Agg</i>	$\mu = -7.66, \sigma = 1.02$
LH-RW	$\mu = -8.66, \sigma = 0.64$	<i>Agg-mpl</i>	$\mu = -1.02, \sigma = 0.87$

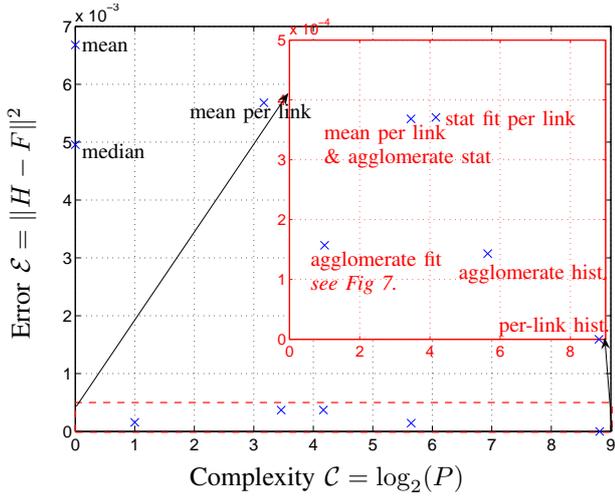


Fig. 5. Comparison of error  $\mathcal{E}$  and complexity  $\mathcal{C}$  for the dynamic walking data. Note inset is the zoomed-in lower portion of the graph

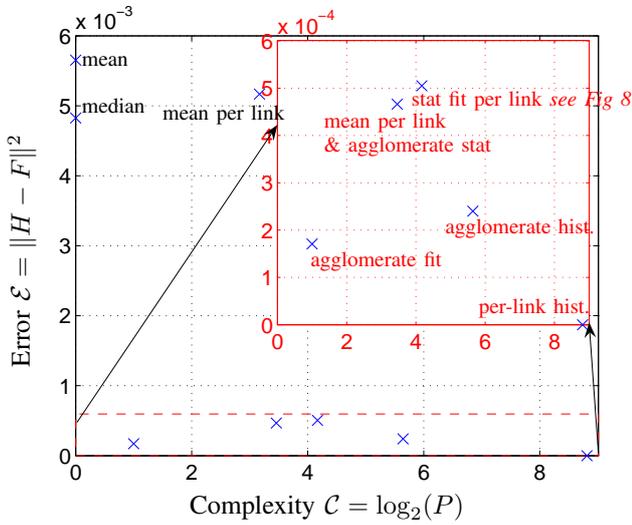


Fig. 6. Comparison of error  $\mathcal{E}$  and complexity  $\mathcal{C}$  for the dynamic running data

relative inaccuracy of mean-per-link (with  $P = M = 9$  parameters), and single mean and median, we once again expand lower portion of the graph to better show the relative error of other model options.

Figs. 5 and 6 show that the lognormal fit to agglomerate data (which is the best agglomerate fit in both cases); is the preferable model in terms of equivalent goodness  $\mathcal{G}$ , performing best in terms of the trade-off between error  $\mathcal{E}$ , and complexity  $\mathcal{C}$ . As in [19] with mean removed from each data set, the Weibull distribution is best fit to the agglomerate data. The Weibull distribution is also the most common best fit for individual per-link data-sets. However it is also not always the best fit, hence there are some different models for different individual link data sets. In contrast to Fig. 2 for the everyday data, the dynamic channel data shows relatively worse error performance in Figs. 5 and 6 for a distribution

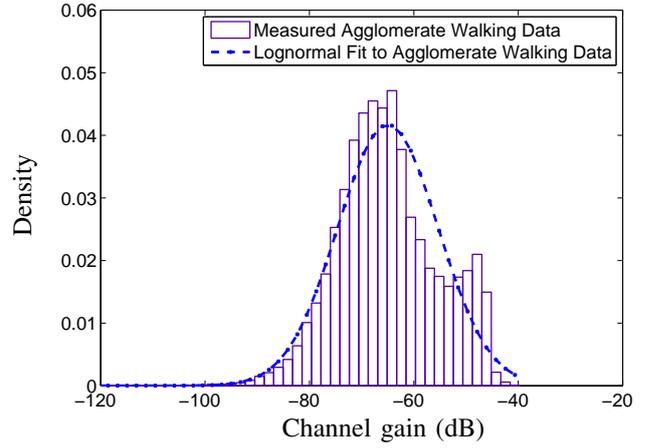


Fig. 7. Probability density, (empirical histogram) with best (Akaike) matched continuous distribution for agglomerate dynamic walking data.

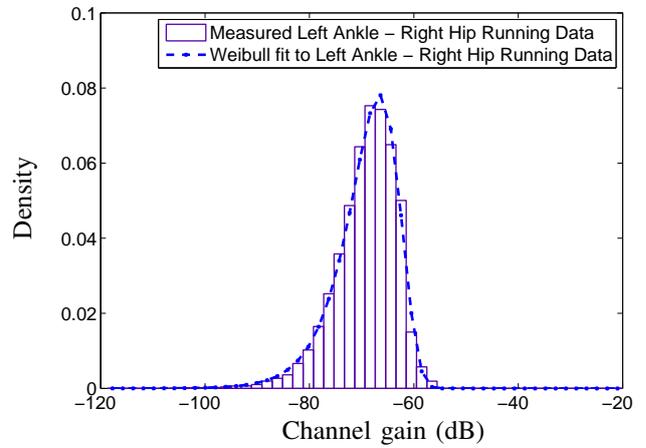


Fig. 8. Probability density, (empirical histogram) with best (Akaike) matched continuous distribution for an example single-link dynamic running data.

fitted to agglomerate data with mean-removed from each link, i.e. for the dynamic data the “mean per link&agglomerate stat” has a worse error performance,  $\mathcal{E}$ , and a significant deterioration in equivalent goodness  $\mathcal{G}$  when compared to the everyday data. Interestingly in Fig. 6 the error performance for the dynamic running data of the agglomerate fit, with respect to all per-link histograms, is better than that of the agglomerate histogram. It can also be seen that Figs. 5 and 6 show similar orders of relative error magnitude, with dynamic data, when compared with Fig. 2 with everyday data.

The statistical lognormal fits for agglomerate data, without mean-per-link removed, and the Weibull fits for agglomerate data with mean-per-link removed, are given in Table II. The Lognormal fit to agglomerate walking data is shown in Fig. 7, where once again the fit is not perfect since the means of each data set are different. An example per-link fit for Left-Ankle to Right-Hip for a subject running is shown in Fig. 8. It should be noted that for the Dynamic data not all fits are this accurate.

TABLE II

BEST LOGNORMAL (LN) FITS TO AGGLOMERATE DYNAMIC DATA; PARAMETERS:  $\mu$  IS LOG-MEAN,  $\sigma$  IS LOG-STANDARD DEVIATION. GIVEN ALSO ARE BEST WEIBULL (W) FITS TO AGGLOMERATE DYNAMIC DATA WITH MEAN-REMOVED FROM EACH PER-LINK SET; PARAMETERS  $a$  IS THE SCALE PARAMETER,  $b$  IS THE SHAPE PARAMETER. *Agg* - AGGLOMERATE DATA; *-mpl* - MEAN-PER-LINK REMOVED

Activity	LN parameters, <i>Agg</i>	W parameters, <i>Agg-mpl</i>
Walking	$\mu = -7.48, \sigma = 1.11$	$a = 0.96, b = 1.68$
Running	$\mu = -6.99, \sigma = 1.41$	$a = 0.89, b = 1.38$

## V. CONCLUDING REMARKS

We presented a goodness-of-fit that combines modeling error (via a metric) and complexity (via log of number of parameters  $P$ ). The function accommodates multiple models and extends the AIC. Best model options can be traded-off, and the method demonstrates different implications for mixed “everyday” data, and set-activity “dynamic” data, but in both cases utility of agglomerate fits (with no normalization) is demonstrated. In all cases serious deficiencies, when considering model error, in choosing a mean (or median) path loss as a sole-characterization of the BAN channel have been demonstrated. Formalization (in future publications) will also consider:

- How to address moderate and large number of samples  $n$ ?
- What if model  $A$  is optimal for AIC (or BIC) while model  $B \neq A$  is optimal for  $\mathcal{G}$  in (4)?
- How are complex analytic models (with few free parameters) impacted?

Our new approach is directly relevant for many types of radio communication (beyond BAN), and we believe can provide new insight to choice of channel model by the radio propagation community.

## REFERENCES

- [1] K. Y. Yazdandoost and K. Sayrafian-Pour, “Channel model for body area network (BAN) ID:-802.15-08-0033,” IEEE submission, Nov. 2008.
- [2] A. Fort, C. Desset, P. De Doncker, P. Wambacq, and L. Van Biesen, “An ultra-wideband body area propagation channel model – from statistics to implementation,” *IEEE Trans. Microwave Theory Tech.*, vol. 54, no. 4, pp. 1820–1826, Jun. 2006.
- [3] W. G. Scanlon and N. E. Evans, “Numerical analysis of bodyworn UHF antenna systems,” *IEEE Elec. Comm. Eng.*, vol. 13, pp. 53–64, Apr. 2001.
- [4] P. S. Hall, Y. Hao, Y. I. Nechayev, A. Alomain, C. C. Constantinou, C. Parini, M. R. Kamarudin, T. Z. Salim, D. T. M. Heel, R. Dubrovka, A. S. Owadall, W. Song, A. Serra, P. Nepa, M. Gallo, and M. Bozzetti, “Antennas and propagation for on-body communication systems,” *IEEE Antennas Propagat. Mag.*, vol. 49, no. 3, pp. 41–58, June 2007.
- [5] P. S. Hall, Y. Hao, and S. L. Cotton, “Advances in antennas and propagation for body centric wireless communications,” in *Europe. Conf. Ant. Prop.*, Apr. 2010.
- [6] R. D’Errico and L. Ouvry, “A statistical model for on-body dynamic channels,” *Springer Intl. J. Wireless Inform. Net.*, Sep. 2010.
- [7] T. Zasowski, G. Meyer, F. Althaus, and A. Wittneben, “Propagation effects in ubw body area networks,” in *Ultra-Wideband, 2005. ICU 2005. 2005 IEEE International Conference on*, September 2005, pp. 16–21.
- [8] D. Neirynek., C. Williams, A. Nix, and M. Beach, “Wideband channel characterisation for body and personal area networks,” in *2nd International Workshop on Wearable and Implantable Body Sensor Networks*, April 2004.
- [9] B. Zhen, M. Patel, S. Lee, E. Won, and A. Astrin, “TG6 technical requirements document (TRD) ID: 802.15-08-0644,” IEEE submission, Sep. 2008.
- [10] D. B. Smith, L. W. Hanlen, J. A. Zhang, D. Miniutti, D. Rodda, and B. Gilbert, “First and second-order statistical characterizations of the dynamic body-area propagation channel of various bandwidths,” *Annals of Telecommunications*, Dec. 2010.
- [11] D. B. Smith, L. W. Hanlen, D. Miniutti, J. A. Zhang, D. Rodda, and B. Gilbert, “Statistical characterization of the dynamic narrowband body area channel,” in *Intl. Symp. App. Sci. Bio-Med. Comm. Tech.*, Aalborg, Denmark, Oct. 2008.
- [12] L. W. Hanlen, D. Miniutti, D. B. Smith, D. Rodda, and B. Gilbert, “Co-Channel interference in body area networks with indoor measurements at 2.4GHz: Distance-to-interferer is a poor estimate of received interference power,” *Springer Intl. J. Wireless Inform. Net.*, vol. 17, Aug. 2010.
- [13] H. Akaike, “A new look at the statistical model identification,” *IEEE Trans. Automat. Contr.*, vol. 19, pp. 716–723, Dec. 1974.
- [14] G. Schwarz, “Estimating the dimension of a model,” *Annals of Statistics*, vol. 6, pp. 461–464, 1978.
- [15] J. Rissanen, “Modeling by shortest data description,” *Automatica*, vol. 14, pp. 465–471, 1978.
- [16] M. Wax and T. Kailath, “Detection of signals by information theoretic criteria,” *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 33, no. 2, pp. 387 – 392, Apr. 1985.
- [17] R. V. L. Hartley, “Transmission of information,” in *International Congress of Telegraphy and Telephony*, Sep. 1927, reprinted in J. Bell Systems, 1928, pp. 535.
- [18] L. W. Hanlen, V. G. Chaganti, B. Gilbert, D. Rodda, T. A. Lamahewa, and D. B. Smith, “Open-source testbed for body area networks: 200 sample/sec, 12 hrs continuous measurement,” in *IEEE PIMRC*, Sep. 2010.
- [19] D. B. Smith, J. A. Zhang, L. W. Hanlen, D. Miniutti, D. Rodda, and B. Gilbert, “Temporal correlation of the dynamic on-body area radio channel,” *IEE Elect. Lett.*, Nov. 2009.