# On the Application of the Baum-Welch Algorithm for Modeling the Land Mobile Satellite Channel

Balázs Matuz, Francisco Lázaro Blasco and Gianluigi Liva

*Abstract*—Accurate channel models are of high importance for the design of upcoming mobile satellite systems. Nowadays most of the models for the land mobile satellite channel (LMSC) are based on Markov chains and rely on measurement data, rather than on pure theoretical considerations. A key problem lies in the determination of the model parameters out of the observed data. In this work we face the issue of state identification of the underlying Markov model whose model parameters are a priori unknown. This can be seen as a hidden Markov model (HMM) problem. For finding the maximum likelihood (ML) estimates of such model parameters the Baum-Welch (BW) algorithm is adapted to the context of channel modeling. Numerical results on test data sequences reveal the capabilities of the proposed algorithm. Results on real measurement data are finally presented.

## I. INTRODUCTION

Satellite services to mobile users are experiencing a renewed interest thanks to the licenses granted for S-band usage for broadcast and interactive services [1]–[3]. The underlying communication channel, referred to as land mobile satellite channel (LMSC), is characterized by strong variations of the received signal power. Obstacles in the propagation path between the satellite and the mobile terminal, such as buildings or trees may cause shadowing or even a complete blockage of the signal. With increasing frequency and decreasing elevation angle such events become more and more likely and strongly impact service availability. A further source of fading is due to multipath propagation: objects in the vicinity of the receiver are source of reflections that cause constructive or destructive interference. In the past several authors proposed Markov chain models to describe the behavior of the LMSC [4]–[8]. The modeling approach can be divided into two stages. First a Markov chain is set up to model slow transitions between different signal levels due to macroscopic effects such as blockage, shadowing, etc. In practice models with two or three states are common, but also a larger number of states is possible. Second, fast signal variations within each state due to multipath are taken into account assuming that the signal amplitude follows some specific distribution. To give an example, a Ricean distribution may be used to describe the signal amplitude in line of sight (LoS) conditions, whereas the amplitude in a blockage state could be assumed to be Rayleigh distributed.

Knowing the underlying channel model, a major issue consists in how to determine the model parameters out of

a sequence of measurement data. In literature there exist several approaches, most of them being rather simple and empirical. In [6] the authors propose first to associate with each measurement sample a state of the underlying Markov chain. This association is done manually. Then, for each state the distribution of the associated samples is approximated by some known distribution by means of curve fitting. In [4] the weighted sum of some known distributions is fitted to the probability density function (p.d.f.) of the measured data. This gives the parameters for the distributions in the different states. Then, each sample is associated with a state by placing thresholds on the signal level. The thresholds are put according to the state probabilities from the fitting step. For highly overlapping distributions, this only works with limited accuracy, as we will show later. A more rigorous attempt is the technique in [7] based on reversible jump Monte Carlo computation [9]. It suggests fully blind estimation, making no prior assumptions on the number of states, nor on the specific distributions, allowing huge flexibility. However, this has the price of a significant increase in complexity and the resulting states and distributions often lack sufficient explanations in terms of underlying physical effects.

Within this work we propose a further way to estimate the model parameters. It exploits the fact that the state identification can be seen as a hidden Markov model (HMM) problem: out of the channel observation we would like to draw conclusions about the underlying Markov process that is not directly observable. A solution to this problem is given by the Baum-Welch (BW) algorithm, that has been widely used in other fields, such as speech or pattern recognition. An application to models of digital channels has already been provided in [10]. In the sequel our focus is on the LMSC. We impose some constraints on the BW algorithm in order to improve its convergence and for sake of simplification. In particular it is well-known that its convergence properties depend on the initial model assumptions. Hence, unlike in [7] we assume prior knowledge on the type of distributions and the distribution parameters (to be obtained by a preceding curve fitting step). Also, we fix the number of states in advance.

The remaining part of the paper is organized as follows. In Section II we recap the BW algorithm. Further we introduce a log-domain computation of the forward-backward metric of the BW algorithm and discuss some adaptations. Section III reports the performance of the algorithm on test data, as well as on sequences of real measurement data. A comparison with the method in [4] is provided.

Balázs Matuz, Francisco Lázaro Blasco and Gianluigi Liva are with the Institute of Communications and Navigation, German Aerospace Center (DLR), Oberpfaffenhofen, 82234 Wessling, Germany. Email: {Balazs.Matuz, Francisco.LazaroBlasco, Gianluigi.Liva}@dlr.de.

## II. OVERVIEW OF THE BW ALGORITHM

Following the footsteps of [11], we consider next the problem of associating an observed sample with a state of our HMM. The BW algorithm can be applied to maximize the probability of a state given the entire observation sequence. Let us denote as $X_t$ the state of the HMM at time $t$, and as $\mathbf{r} = (r_1, r_2, \ldots, r_n)$ the vector of $n$ observations. The problem can be formalized as follows: given the vector of $\mathbf{r}$, we are interested in locally calculating the probability of being in state $i$ at time $t$, i.e. $\Pr\{X_t = i | \mathbf{r}\}$. We define

$$g_t(i) \triangleq \Pr\{X_t = i | \mathbf{r}\} = \frac{f_{\mathbf{R}, X_t}(\mathbf{r}, X_t = i)}{f_{\mathbf{R}}(\mathbf{r})}. \tag{1}$$

For now, we focus on the joint p.d.f. in the enumerator. Under the HMM assumption, after dropping the subscripts for simplicity, this p.d.f. can be rewritten as

$$f(\mathbf{r}, X_t = i) = \underbrace{f(\mathbf{r}_1^t, X_t = i)}_{a_t(i)} \cdot \underbrace{f(\mathbf{r}_{t+1}^n | X_t = i)}_{b_t(i)}. \tag{2}$$

Here we used the shorthand $\mathbf{r}_k^w$ to denote the elements $(r_k, r_{k+1}, \ldots, r_w)$ of the observation sequence $\mathbf{r}$, with $w > k$. Further, referring to (2), we define a forward metric $a_t(i)$ and a backward metric $b_t(i)$. It follows that

$$g_t(i) = \frac{a_t(i) b_t(i)}{\sum_{i=1}^m a_t(i) b_t(i)}, \tag{3}$$

where the normalization by $\sum_{i=1}^m a_t(i) b_t(i)$ corresponds to $f_{\mathbf{R}}(\mathbf{r})$ in (1) and $m$ denotes the number of states. Being $p_{ij}$ the transition probability from state $i$ to $j$, $p_i$ the probability of state $i$ and $f_i(r)$ the probability density function given state $i$, the forward and the backward metric can be computed iteratively as

$$a_t(i) = f_i(r_t) \sum_{j=1}^m a_{t-1}(j) p_{ji}, \tag{4}$$

$$b_t(i) = \sum_{j=1}^m b_{t+1}(j) p_{ij} f_j(r_{t+1}), \tag{5}$$

with the initial metrics $a_1(i) = p_i f_i(r_1)$ and $b_n(i) = 1 \forall i$.

Figure 1 shows an excerpt of a state diagram for a Markov chain in the interval $[t-1, t+1]$. The nodes of the trellis diagram at each time instance denote one of the three possible states, whereas the lines denote all possible transitions. Consider for example state 1. As indicated by the solid arrows, the metrics from all the nodes at $t-1$, as well as $t+1$ contribute to the calculation of the probability of state 1 at time $t$. The most likely state sequence can be determined by choosing the state with the highest probability at each time instance.

Further, we may wish to estimate the probability of having a transition from state $i$ at time $t$ to state $j$ at time $t+1$, given the observation $\mathbf{r}$. This can be expressed as

$$z_t(i, j) \triangleq \Pr\{X_t = i, X_{t+1} = j | \mathbf{r}\},$$

and it turns out that

$$z_t(i, j) = \frac{a_t(i) p_{ij} f_j(r_{t+1}) b_{t+1}(j)}{\sum_{i=1}^m \sum_{j=1}^m a_t(i) p_{ij} f_j(r_{t+1}) b_{t+1}(j)}. \tag{6}$$



Fig. 1. Calculation of the forward-backward metric for 3 states (linear domain).

Even if the BW algorithm is in principle more general, we restrict ourselves to the simple case where the density functions $f_i(r)$, $i = 1, \ldots, m$, are perfectly known, whereas we do not have any knowledge about the transition probabilities $p_{ij}$, which we want to estimate. To do so, we chose some initial values for $p_{ij}$,[1] run the forward-backward algorithm and re-estimate the transition probabilities $\hat{p}_{ij}$ and the initial state probabilities $\hat{p}_i$ according to the re-estimation formulae

$$\hat{p}_{ij} = \frac{\sum_{t=1}^{n-1} z_t(i, j)}{\sum_{t=1}^{n-1} \sum_{j=1}^m z_t(i, j)}, \qquad \hat{p}_i = \sum_{j=1}^m z_1(i, j).$$

Former values of $p_{ij}$ and $p_i$ are replaced by the new estimates and the forward-backward algorithm is run again, leading to updated estimates, which are then fed-back. This process is iterated several times. The state identification step corresponds to the E-step of the expectation-maximization (EM) algorithm, where the model parameters are assumed to be fixed. The re-estimation step corresponds to the M-step, where the most likely model parameters are determined given the hidden state sequence.

A final remark is related to the convergence of the algorithm. It is well known that the EM algorithm, as well as its special instance, the BW algorithm, increases the likelihood of the model iteration by iteration till it converges to a maximum value [12]. However, the algorithm may converge to a local maximum of the likelihood function, rather than to a global one. The convergence of the algorithm can be facilitated by limiting the set of a priori unknown model parameters. Alternatively, a set of various starting points can be considered.

### A. Log-domain Implementation of the BW Algorithm

Already for short observation sequences ($n > 100$) the forward-backward metric may get numerically unstable. As

---

[1]In principle the choice is arbitrary. Nevertheless initial values not too far from the real values facilitate the convergence of the BW algorithm. Good starting points can be found in literature (e.g. in [4]).

a solution, for each $t$ a normalization of the metric is usually performed [12]. Alternatively, a log-domain representation of the corresponding equations is proposed here. Let us define log-probabilities as $\gamma_t(i) \triangleq \ln g_t(i)$, $\alpha_t(i) \triangleq \ln a_t(i)$ and $\beta_t(i) \triangleq \ln b_t(i)$, with $\ln(\cdot)$ being the natural logarithm. Then, (3) can be rewritten as

$$\gamma_i(t) = \alpha_i(t) + \beta_i(t) - \ln \sum_{i=1}^{m} \underbrace{\exp(\alpha_i(t) + \beta_i(t))}_{\exp(\kappa_i)}.$$

Note that the last term can be solved in the log-domain by applying recursively the so-called $\max^*$ operator (also known as Jacobi logarithm) that is defined as $\max^*(\kappa_1, \kappa_2) \triangleq \ln(\exp(\kappa_1) + \exp(\kappa_2))$. Exploiting the identity

$$\max{}^*(\kappa_1, \kappa_2) = \max(\kappa_1, \kappa_2) + \ln(1 + \exp(-|\kappa_1 - \kappa_2|))$$

and noticing that $\max^*(\kappa_1, \kappa_2, \kappa_3)$ can be recursively calculated as $\max^*(\kappa_1, \max^*(\kappa_2, \kappa_3))$, we have

$$\gamma_i(t) = \alpha_i(t) + \beta_i(t) - \max_{i=1:m}{}^*(\alpha_i(t) + \beta_i(t)).$$

In a similar manner, using the shorthand $\phi_i(r_t) \triangleq \ln f_i(r_t)$, $\pi_{ij} \triangleq \ln p_{ij}$ and $\pi_i \triangleq \ln p_i$ we have that the recursions

$$\alpha_i(t) = \phi_i(r_t) + \max_{j=1:m}{}^*(\alpha_j(t-1) + \pi_{ji})$$

with $\alpha_i(1) = \pi_i + \phi_i(r_1)$ and

$$\beta_i(t) = \max_{j=1:m}{}^*(\beta_j(t+1) + \pi_{ij} + \phi_j(r_{t+1}))$$

with $\beta_i(n) = 0$, $\forall i$. Finally, for the re-estimation of the BW metrics we define $\zeta_t(i,j) \triangleq \ln z_t(i,j)$. Taking (6) we have

$$\zeta_t(i,j) = \alpha_t(i) + \pi_{ij} + \phi_j(r_{t+1}) + \beta_{t+1}(j) - $$
$$- \max_{i=1:m}{}^*\left(\max_{j=1:m}{}^*(\alpha_t(i) + \pi_{ij} + \phi_j(r_{t+1}) + \beta_{t+1}(j))\right).$$

The estimation of the parameters proceeds as

$$\pi_{ij} = \max_{t=1:n-1}{}^*\zeta_t(i,j) - \max_{t=1:n-1}{}^*\left(\max_{j=1:m}{}^*\zeta_t(i,j)\right)$$

while

$$\pi_i = \max_{j=1:m}{}^*\zeta_1(i,j).$$

### B. Restrictions on the BW Algorithm

For modeling the LMSC applying the BW algorithm some a priori restrictions on the channel parameters have been applied. This is mainly motivated by two reasons. First, if reasonably good estimates of some channel parameters are available, their use may facilitate the convergence of the algorithm. Second, we consider important that the obtained results have a clear physical interpretation. To give an example, we would like states to be associated with different physical events, such as blockage of the signal or direct LoS. During this work the following restrictions have been applied:

- The type of distributions to be used has been fixed in advance. The original BW algorithm allows estimating the densities $f_i(r)$ iteratively as a mixture of Gaussian distributions [11]. It is however well-established that typical propagation conditions (blockage or LoS, for instance) can be accurately modeled by known distributions.



Fig. 2. Re-estimated transition probability $\hat{p}_{12}$ vs. Bhattacharyya distance for the BW algorithm and threshold methods.

- Estimates of the distribution parameters are provided to the BW algorithm. Such estimates can be obtained for instance by a curve fitting step and are kept fixed through the BW re-estimation. Alternatively at each iteration the estimates could be refined, given the intermediate results.
- The number of states is fixed in advance corresponding to some physical events, such as total blockage of the signal by obstacles or LoS .

### III. APPLICATIONS OF THE BW ALGORITHM

The capabilities of the BW algorithm on different data sets are evaluated next. First we generate artificially a test sequence of samples and run iterative re-estimation. Knowing the original model parameters, our goal is to assess the quality of the re-estimations provided by the BW algorithm. A comparison with the commonly used threshold method and some derivatives is done. Second, the BW algorithm is applied to data obtained from a measurement campaign.

#### A. Application on Test Data Sequences

Given a Markov chain with transition probabilities $p_{ij}$ we generate a sequence of states $\mathbf{x} = (x_1, x_2, \ldots, x_n)$. For each state, an observation sample according to the associated p.d.f. is produced. For simplicity, we fix the number of states $m$ to 2. For state 1, we choose a Gaussian distribution with standard deviation $\sigma_1 = 0.2$. To perform different tests, the mean value $\mu_1$ ranges from $0.4$ to $0.9$. The Gaussian distribution associated with state 2 has mean $\mu_2 = 1$ and variance $\sigma_2 = 0.2$. Since typically the LMSC is highly correlated [6], we choose the state transition probabilities of the Markov chain

$$\begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = \begin{bmatrix} 0.950 & 0.050 \\ 0.025 & 0.975 \end{bmatrix},$$

with corresponding state probabilities $p_1 = 0.333$ and $p_2 = 0.667$. The length of the state sequence (observation sequence) was set to $n = 100000$.

Given the observations $\mathbf{r} = (r_1, r_2, \ldots, r_n)$ and the knowledge on the p.d.f.s, our iterative re-estimation algorithm is ran

to determine the state sequence $x_t$, for $t = 1 \ldots n$, as well as the transition probabilities $p_{ij}$ and the state probabilities $p_i$. It should be obvious that the closer the mean values of both Gaussian distributions are, the bigger shall be the deviation between the re-estimated state sequence $\hat{\mathbf{x}} = (\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n)$, the associated re-estimated transition probabilities $\hat{p}_{ij}$, as well as state probabilities $\hat{p}_i$ and the actual values. To measure the distance of the two distributions $f_1(r)$, $f_2(r)$ associated with the two states, we make use of the Bhattacharyya distance

$$\mathcal{B}\left(f_1(r), f_2(r)\right) = -\ln \int_{-\infty}^{\infty} \sqrt{f_1(r) \cdot f_2(r)} \; dr.$$

For sake of comparison we also apply the threshold method to separate the states [4]. Samples below the threshold $\tau$ are associated with one state, the ones above with the other. We select the threshold $\tau$, such that the average error probability

$$\overline{p}_e \triangleq p_1 \int_{\tau}^{\infty} f_1(r) \; dr + p_2 \int_{-\infty}^{\tau} f_2(r) \; dr,$$

is minimized. In addition, we assume a priori knowledge of the state probabilities (which could be e.g. provided by a previous curve-fitting step). For two Gaussian distributions with variances $\sigma_1 = \sigma_2 = \sigma$, this yields

$$\tau = \frac{\mu_1 + \mu_2}{2} + \frac{\sigma^2 \ln \frac{p_1}{p_2}}{(\mu_2 - \mu_1)}.$$

Further, to suppress frequent state transitions (crossing of the threshold) we apply moving average filtering on the observation sequence. The span of the moving average is set to 10 or 20 samples.

Figure 2 illustrates the estimated transition probability $\hat{p}_{12}$ versus the Bhattacharyya distance for the BW algorithm and the threshold methods with and w/o filtering. Despite close mean values of both distributions, the BW algorithm provides always accurate estimates for the state transition probability $p_{12}$, whereas the threshold methods typically fail when $\mathcal{B}(f_1(r), f_2(r)) < 0.4$. Empirically we found that best results for the threshold methods can be obtained with an averaging window span of 10 samples. It shall be noted however that in this case the resulting state probabilities deviate remarkably as illustrated in Table I. It turns out that with increasing averaging window size even for $\mathcal{B}(f_1(r), f_2(r)) = 0.78$ the estimated state probability $\hat{p}_1$ is too low. Figure 3 depicts the share of wrongly labeled states in the estimated state sequence $\hat{\mathbf{x}}$, obtained through a comparison of the original state sequence $\mathbf{x}$ with $\hat{\mathbf{x}}$. Again the BW algorithm provides by far the best results, followed by the threshold methods with filtering. Note that for low Bhattacharyya distances the share of errors converges to 0.33 which corresponds to $p_1$.

## B. Application on Measurement Data

In fall 2008 a vast measurement campaign was carried out along the US East Coast in the framework of the European Space Agency (ESA) funded MiLADY project [13]. During the field trials the signal levels of the four satellite digital audio radio service (SDARS) satellites were recorded with a mobile vehicular receiver. A statistical channel model was derived out of the collected measurement data employing the BW

TABLE I
$\hat{p}_1$ FOR BW AND VARIOUS THRESHOLD METHODS: UNFILTERED (T1), FILTERED WITH WINDOW SIZE 10 (T10) AND 20 (T20) SAMPLES.

| $\mathcal{B}$ | $\mu_2 - \mu_1$ | BW | T1 | T10 | T20 |
|---|---|---|---|---|---|
| 1.13 | 0.60 | 0.33 | 0.33 | 0.31 | 0.29 |
| 0.78 | 0.50 | 0.33 | 0.32 | 0.30 | 0.28 |
| 0.50 | 0.40 | 0.33 | 0.31 | 0.28 | 0.26 |
| 0.28 | 0.30 | 0.33 | 0.28 | 0.22 | 0.20 |
| 0.13 | 0.20 | 0.33 | 0.22 | 0.07 | 0.04 |
| 0.03 | 0.10 | 0.33 | 0.08 | 0.00 | 0.00 |



Fig. 3. Share of wrongly identified states vs. Bhattacharyya distance for the BW algorithm and threshold methods. A state at time $t$ is considered to be wrongly identified if $\hat{x}_t \neq x_t$.

algorithm. The proceedings are as follows: we first perform a curve-fitting step on the overall p.d.f. of $\mathbf{r}$ similar to [4]. We obtain parameters of the distributions in the different states which serve as input for the BW algorithm. The resulting state probabilities are used to initialize $a_1(t)$. The curve fitting is performed using simulated annealing (SA) [14], a fast meta-heuristic method for global optimization. In case the function to be optimized has several local maxima SA may overcome these and converge to the global minimum. Following literature, we chose three simple distributions to characterize the fast signal variations in the different states. The signal amplitude is assumed to follow a Rice distribution in case of direct LoS to the satellite. We associate a lognormal distribution with the shadowing state and assume that the signal amplitude in the blockage state is Rayleigh distributed. As second step, a preprocessing stage is required. The fast signal variations within a state are known to be correlated (see e.g. [15]). However, (2) implicitly assumes independency among observation samples given a certain state. To comply with the independence assumption, the measurement data is down-sampled, taking into account the coherence time of the process[2]. This leads the final observation $\mathbf{r}$. Finally, given the three distributions and the observed sequence, the BW algorithm is applied as described in Section II.

Let's consider a typical US urban environment with a

[2]The spatial separation between samples after down-sampling was chosen to be 1 $m$ in accordance with [4].

Fig. 4. Curve fit on measurement data for urban environment and a satellite elevation of $30°$.

TABLE II
MEAN STATE DURATION $\bar{D}$ IN METERS AND STATE PROBABILITY $p_i$ FOR THE 3 PROPAGATION STATES.

|  | Method | LoS | Shadowing | Blockage |
|---|---|---|---|---|
| $p_i$ | BW | 0.66 | 0.14 | 0.20 |
|  | T1 | 0.66 | 0.14 | 0.20 |
|  | T10 | 0.70 | 0.12 | 0.18 |
| $\bar{D}_i$ | BW | 22.42 | 4.11 | 26.88 |
|  | T1 | 7.25 | 1.48 | 8.90 |
|  | T10 | 64.35 | 6.87 | 25.15 |

satellite elevation of $30°$. The solid line in Figure 4 shows the p.d.f. of the measured signal envelope, whereas the dashed lines with markers give the results of curve fitting using the three p.d.f.s specified previously. The weighted sum of the Rice, lognormal and Rayleigh distributions is also plotted (dashed with diamonds) and turns out to be close to the p.d.f. of the measured data. The Bhattacharyya distance between the lognormal (Rayleigh) and Rice (lognormal) distribution is 0.5 (1.1), thus posing challenges for state identification. Table II gives the mean state durations $\bar{D}_i = 1/(1 - p_{ii})$ and state probabilities $p_i$ obtained with different state identification methods with minimum state duration set to $1\ m$. Results for the BW algorithm, the threshold method from [4], as well as the modified threshold method with a filter length of 10 samples are shown. It can be observed that the BW algorithm and the threshold method yield the same state probabilities $\hat{p}_i$ as obtained by means of curve fitting. However, the mean state durations obtained by the threshold method are very short. As illustrated in Figure 2 the threshold method tends to over-dimension $\hat{p}_{ij}$, thus to under-dimension $\bar{D}_i$. If a prior filtering step is applied, the state durations become longer than the ones obtained with BW. This is caused by an under-dimensioning of $\hat{p}_{ij}$ for Bhattacharyya distances greater than 0.5 (c.f. Figure 2). Here, the state probabilities are no longer preserved.

## IV. CONCLUSION

This work investigates the application of the BW algorithm to determine the parameters for a LMSC model out of a set of measurement data. The BW algorithm, allows estimating iteratively the hidden state sequence and the transition probabilities of the underlying HMM even for highly overlapping states. Especially in environments with frequent shadowing events conventional methods, such as the threshold methods may lead to inaccurate results on the state transition matrix (STM) of the hidden Markov process. Adaptations of the BW algorithm presented here guarantee numerical stability, as well as proper convergence at manageable complexity. Adaptations to channels different from the LMSC are possible and may be a matter of future investigation.

## V. ACKNOWLEDGEMENTS

## REFERENCES

[1] A. Bolea Alamanac, P. Burzigotti, R. De Gaudenzi, G. Liva, H. N. Pham, and S. Scalise, "In-depth analysis of the satellite component of DVB-SH: Scenarios, system dimensioning, simulations and field trial results," *International Journal of Satellite Communications and Networking*, vol. 27, no. 4-5, pp. 215–240, 2009. [Online]. Available: http://dx.doi.org/10.1002/sat.933

[2] S. Scalise, C. Niebla, G. Gallinaro, M. Andrenacci, R. Rinaldo, O. Del Rio Herrero, M. Breiling, D. Finocchiaro, J. Cebrian Puyuelo, and G. Schlueter, "System design for Pan-European MSS services in S-band," in *Advanced Satellite Multimedia Systems Conference (ASMS) and the 11th Signal Processing for Space Communications Workshop (SPSC), 2010 5th*, 2010, pp. 538 –545.

[3] Satellite Digital Audio Radio Service (SDARS). [Online]. Available: http://www.sirius.com/

[4] E. Lutz, D. Cygan, M. Dippold, F. Dolainsky, and W. Papke, "The land mobile satellite communication channel - recording, statistics, and channel model," in *IEEE Trans. Vehicular Technology*, vol. 40, no. 2, May 1991, pp. 375–386.

[5] Y. Karasawa, K. Kimura, and K. Minamisono, "Analysis of availability improvement in LMSS by means of satellite diversity based on three-state propagation channel model," in *IEEE Trans. Vehicular Technology*, vol. 46, no. 4, November 1997, pp. 957–1000.

[6] F. Perez-Fontan, M. Vazquez-Castro, C. Cabado, J. Garcia, and E. Kubista, "Statistical modeling of the lms channel," in *IEEE Trans. Vehicular Technology*, vol. 50, no. 6, November 2001, pp. 1549–1567.

[7] C. Alasseur, S. Scalise, L. Husson, and H. Ernst, "A novel approach to model the land mobile satellite channel through reversible jump Markov chain Monte Carlo technique," *IEEE Transactions on Wireless Communications*, vol. 7, pp. 532–542, 2008.

[8] S. Scalise, H. Ernst, and G. Harles, "Measurement and modeling of the land mobile satellite channel at Ku-band," *Vehicular Technology, IEEE Transactions on*, vol. 57, no. 2, pp. 693 –703, 2008.

[9] P. Green, "Reversible jump MCMC computation and Bayesian model determination," in *Biometrika*, vol. 82, no. 40, 1995, pp. 711–732.

[10] W. Turin and M. M. Sondhi, "Modeling Error Sources in Digital Channels," *IEEE Journal on Selected Areas in Communications*, vol. 11, pp. 340–347, 1993.

[11] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," in *Proceedings of the IEEE*, February 1989, pp. 77(1):257–286.

[12] W. Turin, *Performance Analysis and Modeling of Digital Transmission Systems (Information Technology: Transmission, Processing and Storage)*, 1st ed.   Springer, May 2004.

[13] Mobile satellite channeL with Angle DiversitY. [Online]. Available: http://telecom.esa.int/telecom/www/object/index.cfm?fobjectid=29020

[14] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by Simulated Annealing," *Science, Number 4598, 13 May 1983*, vol. 220, 4598, pp. 671–680, 1983.

[15] W. C. Jakes, *Microwave mobile communications*.   Wiley, New York, 1974.