

## Bayesian DHI using passive seismic low frequency data

Nima Riahi\*, Mike Kelly, Martine Ruiz, Weiwei Yang, Spectraseis AG

### Summary

We present a procedure for producing a Bayesian DHI for low frequency passive seismic (LFPS) data. The approach utilizes two LFPS attributes to classify and determine the likelihood of hydrocarbon existence in the subsurface. The attributes are based on strength and variability of the empirically observed hydrocarbon tremor. An improved, more robust tremor energy measure based on the temporal characteristics of the signal is presented and used. Bayesian classification is employed both to accommodate uncertainties in the data and to provide a risk estimate.

The process was tested over four fields with known surface projection of the oil-water contact (OWC). Prediction results correlate well with reservoir locations. Accuracy and significance of results will be discussed along with possible extensions. The approach provides a rigorous method for producing hydrocarbon probabilities based on LFPS data.

### Introduction

Ambient seismic noise recordings made with broadband seismometers over hydrocarbon-bearing structures have been observed to exhibit spectral anomalies in the frequency range of about 1 to 6 Hz, most notably in the vertical displacement (Dangel et al., 2003; van Mastrigt and Al-Dulajjan, 2008; Saenger et al., 2009; Lambert et al., 2009). Although these anomalies have been found to be associated with hydrocarbons, there is no accepted theory for the generating mechanism. A preliminary theory states that the anomaly is a microtremor associated with the presence of a multiphase fluid: brine and hydrocarbons (Saenger et al., 2009). The tremor is visible as redirected spectral energy over the seismic background.

The microtremor strength was shown to be indicative of hydrocarbon (HC) presence in the subsurface (Saenger et al., 2009; Lambert et al., 2008). Both short-term (e.g. transient car noise) and long-term (e.g. daytime anthropogenic activities) interferences in the LF band require a careful editing of the time intervals over which the recording is analyzed. Using the quietest time windows the tremor energy is calculated by summing the short-time power spectral densities (PSDs). The energy is used as a low frequency (LF) attribute plotted on a map or in profile. Even when using minimum noise-affected data, the LF attributes often exhibit unexpected variability and the resulting maps are difficult to interpret. This is likely a consequence of using a sum of PSDs of a non-stationary signal.

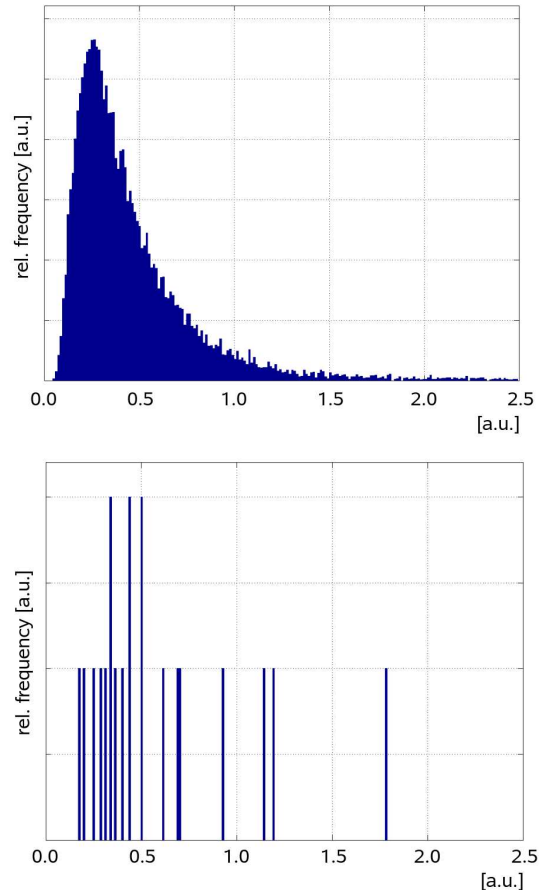


Figure 1: Spectral energy in the frequency range 1-4 Hz of a 100-hour broadband recording was sampled using a sampling window size of 40 sec. The resulting distribution in time is shown on top. When using a 3-hour sampling window size the distribution in the bottom results. The overall characteristics of the distributions are very similar for both sampling window lengths.

### Description of the new prediction process

Empirical evidence and forward modeling experiments suggest that tremor variability can also be indicative of HC presence (van Mastrigt and Al-Dulajjan, 2008; Lambert et al., 2009). The statistical process presented here for LFPS data employs two attributes: tremor strength and variability. Both attributes are based on the tremor energy distribution

## Bayesian DHI using LF passive seismic data

in time. We will show that distribution based attributes are more robust compared with the previously used methodology.

A Bayesian HC probability is then computed based on two groups of HC and no HC exemplar receivers. These exemplars are chosen based on the characteristics of their microtremor signature and, if available, prior information about the field (e.g. logs from nearby wells).

The resulting probability maps can be readily used for risk assessment. Furthermore, they are also more accurate when compared to a conventional single-attribute, deterministic classification process. The details of the process are given in the following three sections.

### Calculation of strength and variability

Tremor energy is calculated by selecting a time window (a sample) in a recording, computing its power spectral density (PSD), and then integrating it over the LF band of interest. The time evolution of this energy was systematically studied on passive data both over and away from HC. Figure 1 shows the distribution in time of the energy for a long recording receiver with 100 hours of quiet recording time. Two sample window sizes were used: 40 seconds (top) and 3 hours (bottom). We see that the underlying distribution is lognormal and that its characteristics are largely invariant to the sample window size.

We conclude that at least part of the variability of conventional tremor energy maps is due to chance errors brought about by using just one large window on a non-stationary ambient background signal. Note that many of the 3 hour samples in Figure 1 (bottom) would have overestimated the likely tremor energy. Statisticians call this the *sampling error*. We therefore sampled the tremor energy with small time windows to resolve its distribution characteristics. The median  $m$  of the distribution is then used as a robust energy measure. Also, the sample standard deviation  $\sigma$  can be calculated as a measure of tremor variability during the considered time windows.

The above observations largely hold as well for other energy attributes as used by Saenger et al. (2009).

### HC discrimination in $m$ - $\sigma$ space

Figure 2 shows histograms of LF energy median  $m$  and standard deviation  $\sigma$  from 87 receivers from a field in West Texas (carbonate oil reservoir at a depth of ~2100 m). These spatial distributions are also of lognormal nature, an observation supported by our research over a substantial number of other fields. In Figure 3, the 87 values for  $m$  and

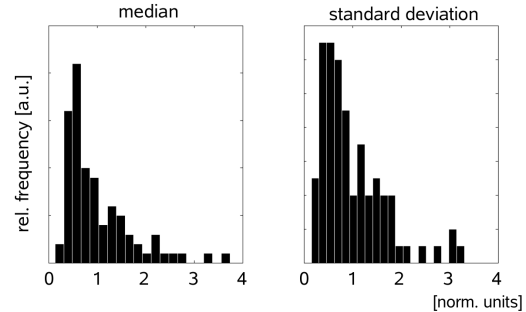


Figure 2: Distribution of tremor median  $m$  (left) and standard deviation  $\sigma$  (right) over a field in West Texas (normalized values used). Both distributions are approximately lognormal.

$\sigma$  are plotted against each other. Each axis was scaled so that the mean of its values equals 1. The known surface projection of the OWC was used to mark the circles as HC (filled circle) or no HC (open circles). Note that the OWC information in Figure 3 is *not* part of the process but merely used to motivate the use of the  $m$ - $\sigma$  space for HC discrimination. It is apparent that in this space the receivers separate to a high degree into two groups, based on whether the recording was made over hydrocarbons or not.

The separation is, however, not perfect. Most likely causes for this include:

- Noise sources near a receiver which alter its LF energy (Nguyen et al., 2009). Time windowing cannot entirely exclude this.
- Site effects on the LF wave field caused by the overburden (Bard 1999; Okada 2003).
- Interference among tremors from distributed HC patches in the subsurface and between surface noises (Lambert et al., 2009).

Furthermore, different reservoirs might exhibit different tremor signatures as these are likely to vary depending on fluid type and rock properties.

### Bayes probability calculation

The above mentioned uncertainties in the data make an empirical Bayesian methodology a natural choice for data classification. First, its outputs are probabilities – a prerequisite for uncertainty-aware risk analysis. Second, its empirical element allows an interpreter to include prior information into the prediction process. This information can be, e.g. log data from nearby wells, or partially known

## Bayesian DHI using LF passive seismic data

reservoir data (exploitation). Figure 4 shows the resulting  $m$ - $\sigma$  space after two groups of receivers were marked: red circles are receivers chosen as HC exemplars and blue circles are no HC exemplar receivers. The process was tested in a blind fashion, i.e. *only tremor signature and no prior information was considered for selecting exemplar receivers*.

To calculate hydrocarbon probabilities using the Bayesian methodology, so-called class-conditional probability density functions (pdf's), or *exemplar pdf's*, are required (see e.g. Carlin, 2000). These exemplar distributions will now be estimated using the selected exemplar data groups. Since both strength and variability are seen to be roughly lognormally distributed (Figure 2), a parametric lognormal model was selected for the exemplar pdf's. Model parameters are calculated from the  $(m, \sigma)$  values of the exemplar receivers by use of Maximum Likelihood estimation. The red and blue contour lines in Figure 4 indicate the lognormal HC and no HC exemplar pdf's based on the marked data points.

A useful feature of the Bayes approach is the concept of prior probabilities. For an HC classification this means that assumed general prospectivity can readily be built into the analysis. In a blind case (such as the one presented here) a so called *uninformative prior probability distribution* is used for all receivers:  $P(\text{HC}) = P(\text{NHC}) = 0.5$ .

Comparing the  $(m, \sigma)$  values of a receiver against the HC/no HC exemplar pdf's will then yield an HC probability for the receiver location. This is done for the

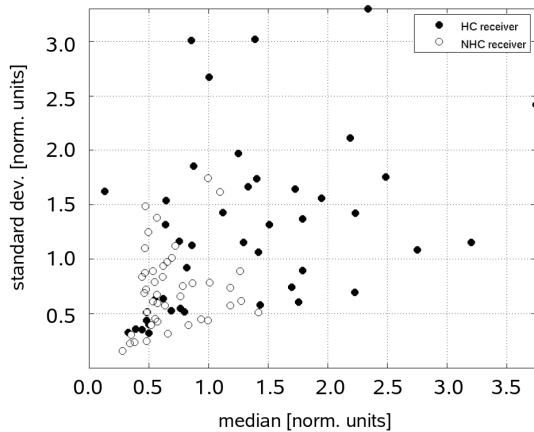


Figure 3: For 87 receivers, microtremor energy distribution median  $m$  and standard deviation  $\sigma$  were plotted against each other. Solid circles and open circles mark receivers above and away from HC, respectively.  $m$  and  $\sigma$  were scaled to have mean = 1. The two groups separate reasonably well.

West Texas dataset and the resulting HC probability map is discussed in the next section.

## Results

Figure 5 shows the kriged HC probability map computed for the West Texas field. The black contour line represents the oil-water contact boundary. The eastern part of the field is sealed off from the reservoir by a fault and is non-prospective – a feature that was well represented by the prediction process. Hydrocarbons are accumulated in two anticlines to the north and south of the narrow neck of the OWC. The process mapped the northern part well but failed to recognize the southern part. Possible reasons for this include: site effects, destructive tremor interference due to spatial distribution of HC patches, or change in reservoir properties affecting the tremor generation mechanism.

The process is blindly tested on four fields. All receivers with a HC probability of  $P(\text{HC}) > 0.5$  are considered HC predictions. Knowledge about surface projected HC locations was used after that to capture the accuracy of the predictions using these measures:

$$\text{HC prediction accuracy } A_{\text{HC}} = \frac{n_{\text{HC}}}{N_{\text{HC}}} ,$$

$$\text{NHC prediction accuracy } A_{\text{NHC}} = \frac{n_{\text{NHC}}}{N_{\text{NHC}}} ,$$

where  $n_{\text{HC}}$  and  $n_{\text{NHC}}$  are the numbers of correct HC/no HC predictions and  $N_{\text{HC}}$  and  $N_{\text{NHC}}$  are the numbers of receivers actually above/away from HC. Table 1 shows those accuracies for the four fields. Because the performance was

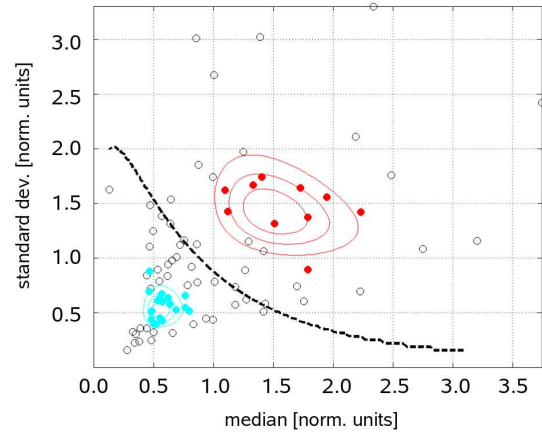


Figure 4: The  $m$ - $\sigma$  space already shown in Figure 3. Based on observed tremor signature two groups of receivers were marked by an interpreter as HC (red circles) and NHC (blue circles). The contours represent the lognormal exemplar distributions estimated from the marked HC/NHC receivers. The dashed line indicates the locus of equal probability of HC/NHC (decision boundary).

## Bayesian DHI using LF passive seismic data

established on a relatively small number of locations per field, the significance of the accuracy must be verified. Therefore, accuracies were computed above which the predictor can be assumed to be non-random with a confidence of >90%. These lower accuracy thresholds are given in Table 1 as well. If a process has accuracies greater than these thresholds it can be considered significantly non-random. The Bayes HC and no HC classifications in fact are greater than these thresholds in most of the cases and we conclude that the process is non-random.

The accuracies of the classification process were compared against accuracies from a deterministic classification that uses natural breaks (using Jenks optimization) in the median values only (Jenks, 1971). Inspection of Table 1 shows that the two-attribute Bayes predictor outperforms the one-attribute predictor in most cases, especially for HC predictions. The process can thus be considered superior to the deterministic one-attribute predictor.

### Conclusions

A two-attribute Bayesian DHI process based on strength and variability of the empirically observed hydrocarbon microtremor has been designed. The process utilizes median and standard deviation of the distribution in time of an LF energy attribute. It was shown that the distribution median offers a more robust energy measure compared to previous techniques that estimated energy based on one large time period. The process gives quantitative HC probability maps that are easy to interpret and can be used for risk analysis.

The process was tested on four fields with known surface-projected HC locations. It was established that the prediction accuracy is significantly (>90%) non-random in most cases. Also, the multi-attribute prediction process was shown to be superior to a deterministic predictor based on tremor energy alone.

Ground noise and subsurface effects can deteriorate the accuracy of the results. Possibilities to improve accuracy include increasing the dimension of the attribute space through more LF attributes (other LF attributes have been

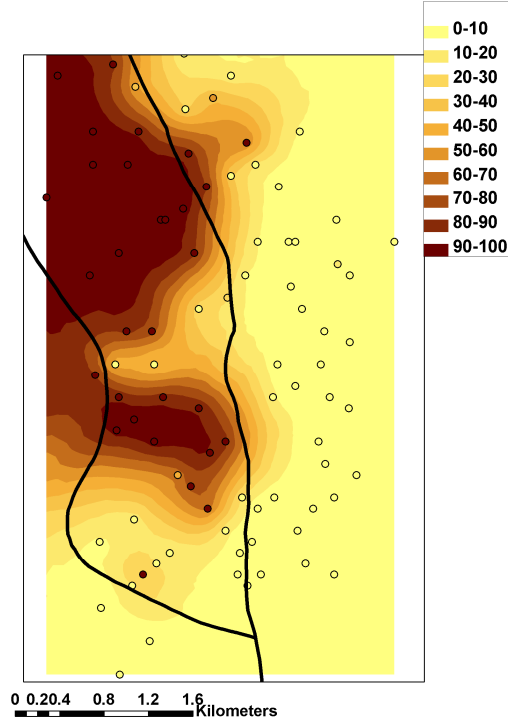


Figure 5: Kriged hydrocarbon probability map from the presented process. The legend indicates HC probability in percent. The black contour line is the surface projected oil-water-contact (OWC) of the reservoir.

suggested by Saenger et al., 2009, and Lambert et al. 2009), or noise filtering in time and space-time (Nguyen et al., 2009). Forward modeling can also help understand the influence of site effects caused by the overburden.

### Acknowledgments

We thank Alex Goertz, Brad Artman, and Rob Habiger for valuable comments and discussions and Konrad Cieslik for support in preparing the plots.

	#HC rec.	#NHC rec.	$A_{HC}$ Jenks	$A_{NHC}$ Jenks	$A_{HC}$ Bayes	$A_{NHC}$ Bayes	min $A_{HC}$ random	min $A_{NHC}$ random
<b>Field 1 (shown)</b>	43	44	56 %	80 %	70 %	91 %	62 %	59 %
<b>Field 2</b>	24	9	67 %	100 %	63 %	89 %	88 %	47 %
<b>Field 3</b>	9	22	67 %	73 %	100 %	77 %	52 %	85 %
<b>Field 4</b>	14	14	29 %	71 %	64 %	79 %	70 %	69 %

Table 1: Accuracies for both HC and NHC prediction are given for the one-attribute deterministic prediction (“Jenks”) and the two-attribute Bayes DHI (“Bayes”). A process is significantly (>90%) non-random if its accuracy is greater than the value given in the last two columns.