Defining the Transfer Functions of the PCAD Model in North Atlantic Right Whales (*Eubalaena glacialis*) – Retrospective Analyses of Existing Data

Scott D. Kraus, Ph.D.
Vice President of Research
New England Aquarium
Central Wharf
Boston, MA 02110
phone: (617) 973-5457 fax: (617) 973-0242 email: skraus@neaq.org

Rosalind M. Rolland, D.V.M.
Senior Scientist
Research Department
New England Aquarium
Central Wharf
Boston, MA 02110
phone: (617) 973-6587 fax: (617) 973-0242 email: rrolland@neaq.org

Award Number: N000141010614
http://www.neaq.org/conservation_and_research/projects/project_pages/right_whale_research.php
http://www.neaq.org/conservation_and_research/projects/project_pages/right_whale_conservation_medicine_program.php

LONG-TERM GOALS

Anthropogenic noise has been shown to cause both behavioral and physiological changes in marine mammals, but the potential for long-term population level effects is not known. The NRC (2005) Population Consequences of Acoustic Disturbance (PCAD) model provides a framework to trace the consequences of acoustic disturbance through the life history of a marine mammal to its population status. In North Atlantic right whales (*Eubalaena glacialis*), extensive data on hormone levels, health and body condition, and individual life history exists. Our long term goal is to analyze the links between stress and thyroid hormones, visual assessments of health, and vital rates of right whales. This research supports the modeling efforts on PCAD transfer functions and develops a theoretical framework for field studies on acoustic disturbance for the North Atlantic right whale.

OBJECTIVES

The first objective is to test an alternative approach to elements of the PCAD model by: 1) substituting "behavior change" with direct measurements of physiological changes (using fecal hormone levels - Hunt et al. 2006; Rolland et al. 2005); 2) replacing "life function" with skin and body condition indices (Pettis et al. 2004), and 3) investigating the links between these parameters and right whale survival, reproduction and maturation (Figure 1). Analyses of retrospective data and new data on fecal thyroid hormones (in FY 2011) will set the stage to apply stress/thyroid hormone data and health indices to assess acoustic disturbance in right whales exposed to different levels of anthropogenic acoustic inputs.

A second objective of this project is to analyze both acoustic recordings and fecal stress hormone levels from the Bay of Fundy in 2001 for the period before and following 9/11 (compared to similar data from other years). We hope to determine if the observed reduction in vessel activity post-9/11 resulted in a decline in ambient noise, and if there was an accompanying decrease in stress hormones in the right whales inhabiting the Bay after this event.

APPROACH

The approach in FY 2010 was to: 1) update the Visual Health Assessment Database by evaluating and scoring all images taken of right whales taken between 2003-2007; 2) Model the relationships between fecal stress hormone levels and the fecal reproductive hormones in order to characterize (and control for) the normal physiological variations in stress hormones that are known to occur with different sexes, ages and reproductive states; 3) Begin modeling the links between body and skin condition scores and fecal stress hormones. (In FY 2011 thyroid hormone values will be added to this model.):4) Analyze acoustic data collected opportunistically before and after 9/11 to compare ambient noise levels before and after that event; and5) Analyze existing fecal stress hormone data collected from 1999-2005 to compare physiologic stress levelsafter 9/11/2001 to fecal stress hormones during all other timeframes.

WORK COMPLETED

Task 1a: Health Assessments (Heather Pettis, Roz Rolland)
The visual health assessment database was updated for all right whale sightings between 2003 and 2007. Details of the visual health assessment method and scoring criteria can be found in Pettis et al. (2004). All sightings of individual whales in a particular habitat were evaluated together (batch) resulting in a single score entry for each assessed parameter for each habitat and year. A total of 4,019 batches, consisting of 114,434 images of 12,262 sightings, were evaluated and scored (Table 1).

Table 1. Total number of batches and sightings evaluated and assessed for visual health assessment database update, 2003-2007.

Year	Batch Count	Sighting Count
2003	584	1,906
2004	564	1,529
2005	913	3,018
2006	832	2,518
2007	1,125	3,289
Total	**4,019**	**12,262**

In addition to updating the visual health assessment database for 2003-2007 sightings, all pre-2003 sightings added to the database or modified since the last update were evaluated. A total of 236 batches consisting of 528 sightings were evaluated for this update. We also re-examined all previously assessed right whale sightings (6,119 batches consisting of 26,744 sightings) for skin condition, lesion assessment and presence.

Tasks 1b and c: Linking Hormone, Health Assessments, and Life History Data (Roz Rolland – hormone data; Scott Kraus – life history data; Peter Corkeron - statistics)
In order to interpret the relationships between hormone levels for differentiating between reproductive classes of right whales, we used a form of classification tree analysis, a relatively recent analytical tool with applications in marine biology and ecology (e.g. De'ath and Fabricius 1999, De'ath 2002); and animal acoustics (e.g. Oswald et al. 2003, Risch et al. 2007). Trees work by recursive binary splitting of multivariate data, performing an exhaustive brute-force search over all possible splits of response (i.e. classifying) variable, and then selecting the best split. Each split results in two mutually exclusive groups (nodes) that are as homogeneous as possible, based on the response variable, and then each smaller (child) node is further split similarly. An issue with the classical formulation of classification trees is when to stop splitting, in order to avoid overfitting the data. This is usually achieved by overfitting, and then pruning trees based on cross-validation (e.g. Venables and Ripley 2002). A recent innovation (Hothorn et al. 2006) has been the development of conditional inference trees, which take a statistical approach to partitioning, accounting for the distributional properties of the data. We used these conditional inference trees (CITs) to explore the relationships between hormone levels and known reproductive state of North Atlantic right whales.

We also tried a more standard modeling approach: using a mixed-effect model with cortisol level as the dependent variable, and reproductive hormone levels as the predictor variables for fixed effects, with reproductive state as the random effect. Rather than seeking an explanatory relationship, we used this approach to investigate outliers in the model. Of the three outliers with higher than predicted levels of cortisol (relative to other hormones), one was a female in her first pregnancy, another a severely entangled adult that died shortly after, and the third, an adult male. One pregnant female with strikingly high levels of reproductive hormones stood out as the only outlier with low cortisol levels. Although less informative than the CIT model, this also demonstrates that hormones sampled form feces can be used to determine animals that are unusually stressed.

Task 2: The 9/11 analysis (Susan Parks – acoustics; Roz Rolland – Hormone data; Peter Corkeron, statistics)
Acoustic data were collected in August and September 2001 for a project related to right whale social behavior. Data were collected with an HTI-94-SSQ hydrophone with a built-in preamplifier (-170 dB re 1V/uPa) recorded into a TASCAM DA-P1 DAT recorder (flat frequency response of the system 50 Hz – 20 kHz ± 1 dB). The signals were bandpass filtered between 50 Hz to 500 Hz prior to measurement. Calibration value for the hydrophone and gain from the recording/acquisition system is -161 dBV/uPa. Records from each day were compiled, and converted to .wav files with original sampling (44.1 kHz). There records were then bandpass filtered between 50 Hz -500 Hz (in Adobe Audition, settings: Blackman 1024 FFT, -20 dB below and above 50 Hz and 500 Hz.) All extraneous noise (splashing/voices/whale sounds/close vessels/etc.) were removed to select the quietest section of background noise in each recording. These final files were then run through a custom Matlab program to calculate SPL dB rms re 1uPa (Table 3). Uncalibrated relative measurements on independently cleaned files were conducted by Manolo Castellote and are reported in Table 2.

There are relatively few (12) sets of hormone samples available from 12th September 2001 to the end of the 2001 field season in the Bay of Fundy [hereafter, post-9/11 data]. Of these, only half had reproductive state information available for them, (and all were from juveniles) so we analyzed the data ignoring reproductive state. The full data set excluding the post-9/11 data [hereafter, non-9/11 data] comprises 247 samples. We developed a model from the non-9/11 data to establish what is "normal" for North Atlantic Right whales, and compared that with the post-9/11 data. This is not the

standard hypothesis-testing approach, where the null hypothesis is that there is no difference between the post-9/11, and the non-9/11 data. Instead we asked whether a model developed on the non-9/11 data adequately describes the post-9/11 data. To do this, we ran a standard linear model with cortisol as the dependent variable, and the other hormones as explanatory variables, then used it to predict stress hormone values for the post-9/11 data. We then compared those predicted values with the actual stress hormone levels from the post-9/11 samples.

Table 2. Summary of level measurements from 2 days before and 2 days after 9/11/2001.

Date	Time	Sea State	Average Relative RMS power measures (Castellote)	Total Relative RMS power (Castellote)	SPL rms dB re 1 uPa (50 Hz - 500 Hz	Noise PeakF (HZ)
8/25/2001	19:15:56	?	-47	-44	107	60
8/29/2001		?	-47	-46	107	60
9/12/2001	5 p.m.	1	-54	-52	101	155
9/13/2001	12:30 p.m.	1	-54	-53	101	165

RESULTS

The updated health assessment database now consists of a total of 10,644 batches and 39,350 sightings. This will provide the foundation for evaluating visual assessment methods as a means of predicting vital rates in right whales, and to investigate the relationships between fecal stress (and thyroid) hormone levels and body/skin condition.

The CIT classification of hormone samples directly assigned to individual whales of known reproductive state is shown in Figure 2. With 97 observations (samples assigned to identified whales), all pregnant females, most (22 of 23) mature males, most (20 of 23) lactating females, and most (19 of 20) juvenile females were correctly assigned. Only resting females (no class determined, so all 8 misclassified) and juvenile males (5 of 14 correctly assigned) were problematic, although just over half of the "juvenile" males assigned to "mature" males (4 of 7) could have been mature or maturing (actual age unknown but at least 6 or 8 yrs, and known minimum age of 8).

One result of the CIT was a node that included animals with relatively high levels of cortisol, given their other hormone levels (node 8 in Figure 2). Of these, two were females nursing their first calf, a mature male incorrectly classified as he had low androgen levels, a resting female killed by ship strike, an immature animal that was infected with *Giardia* and *Cryptosporidium*, a resting female in the middle of a six-year calving interval, and an apparently healthy juvenile. These results indicate that (a) as previous work at NEAq has shown, hormones sampled from feces provide a reliable indicator of stress, and (b) conditional inference trees offer a powerful and successful approach to investigating hormone data.

To analyze the relationship between stress levels and body condition scores, we ran a series of nested linear models, initially with cortisol level as the dependent variable, and the other hormone levels nested in reproductive state as the explanatory variables. This model was used as a base model, to which were added, in turn and separately, each health assessment score, also nested by reproductive

state. In each case, the change in AIC indicated that including body condition scores improved the fit of the model to the data (Table 3). In all cases, mature males with condition indices indicative of poor condition had significantly higher cortisol levels than mature males in better condition.

Table 3. AIC scores for linear models that included body condition scores.

Base model (no body condition score):	792.3178
Model including skin score:	776.2064
Model including body fat score:	756.5413
Model including cyamid score:	727.2558
Model including left side rake mark score:	587.2078
Model including right side rake mark score:	738.9334

In the 9/11 analysis both acoustic methods are in good agreement, showing a 6dB decrease in the noise in the recordings made after September 11th. More importantly, the noise spectrum changed dramatically (Figure 3), with reduction of noise below 150 Hz. In the 9/11 hormone analysis, of the 12 post-9/11 data points, the predicted values were larger for 11, and this remained the case when the non-9/11 data were sub-setted to include only Bay of Fundy samples. Were the model a satisfactory descriptor of the post-9/11 data, 6 of the 12 predicted values should have been smaller than the post-9/11 samples. Because of the small sample size, the power (1-beta) to detect a small effect is poor, and even for a large effect, power only approaches 0.5, so the proportions are not significantly different (Fisher's exact test, p=0.06865). However, the predicted values were larger than the actual stress levels plus one predicted standard error in 10 of the 12 cases, suggesting that despite the small sample size, the model developed on the non-911 data is not a good descriptor of the post-911 samples.

IMPACT/APPLICATIONS

The successful use of fecal stress hormones and visual health assessment methods suggest these may be additional tools for researchers studying species where behavior is cryptic and difficult to study, or where long-term data sets are not available.

RELATED PROJECTS

No related projects.

REFERENCES

De'ath, Glenn. 2002. Multivariate regression trees: A new technique for modeling species-environment relationships. Ecology 83:1105–1117.

De'ath, Glenn, and Katharina E. Fabricius. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecology 81:3178–3192.

Hunt KE, Rolland RM, Kraus SD, Wasser SK. 2006. Analysis of fecal glucocorticoids in the North Atlantic right whale (*Eubalaena glacialis*). General and Comparative Endocrinology 148:260-272.

Hothorn, T., Hornik, K., & Zeileis. A. 2006. Unbiased Recursive Partitioning. Journal of Computational and Graphical Statistics. 15(3): 651-674.

National Research Council of the U.S. National Academies (NRC). 2005. Marine Mammal Populations and Ocean Noise: Determining When Ocean Noise Causes Biologically Significant Effects. National Academy Press, Washington, DC.

Oswald, J. N., Barlow, J. and Norris, T. F. 2003. Acoustic identification of nine delphinid species in the Eastern Tropical Pacific Ocean. Marine Mammal Science, 19:20–037.

Pettis HM, Rolland RM, Hamilton PK, Brault S, Knowlton AR, Kraus SD. 2004. Visual health assessment of endangered North Atlantic right whales (*Eubalaena glacialis*) using photographs. Canadian Journal of Zoology 82:8-19.

Risch, D., Clark, C. W., Corkeron, P. J., Elepfandt, A., Kovacs, K.M., Lydersen, C., Stirling, I. & Van Parijs, S. M. 2007. Vocalizations of male bearded seals, Erignathus barbatus: classification and geographical variation. Animal Behaviour, 73, 747-762.

Rolland RM, Hunt KE, Kraus SD, Wasser SK. 2005. Assessing reproductive status of right whales (*Eubalaena glacialis*) using fecal hormone metabolites. General and Comparative Endocrinology 142:308-317.

Venables, W.N., & Ripley, B.D. 2002. Modern applied statistics with S. New York, Springer, 4th ed.

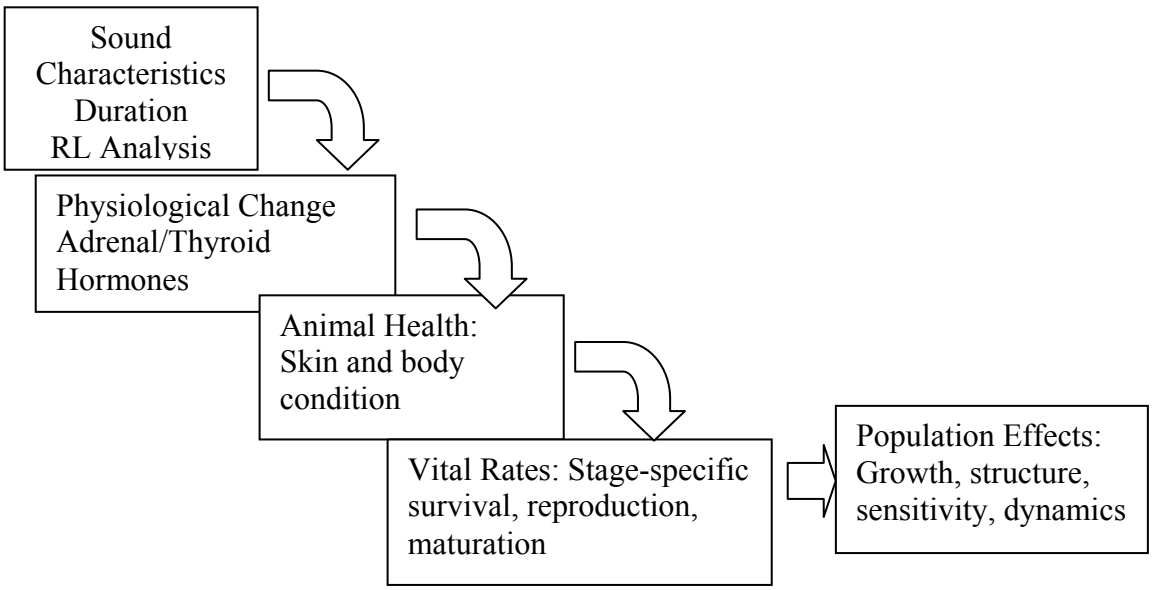

Figure 1. Right whales and acoustic disturbance – the proposed adaptation of the PCAD model. In this alternative framework, we substituted "behavioral changes" resulting from acoustic disturbance with "physiological changes" using adrenal and thryoid hormones, and the "life function changes" with "animal health changes" using visual assessment of body and skin condition.

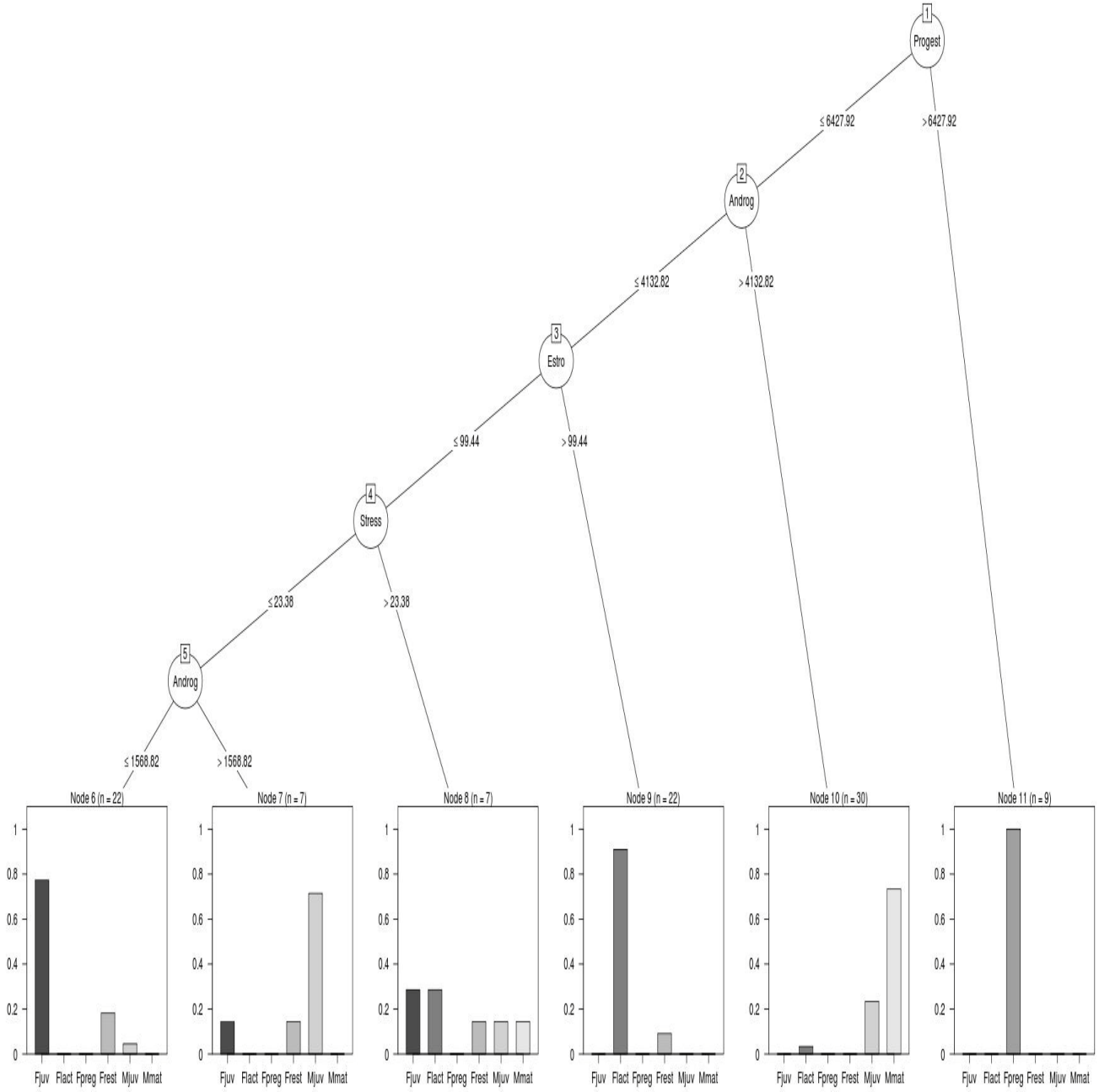

Figure 2. The Conditional Inference Tree of hormone data from samples collected from known individual North Atlantic right whales shows clear separation by age, sex, and reproductive state. The nodes split out pregnant and juvenile females, adult males, and lactating females with high accuracy.

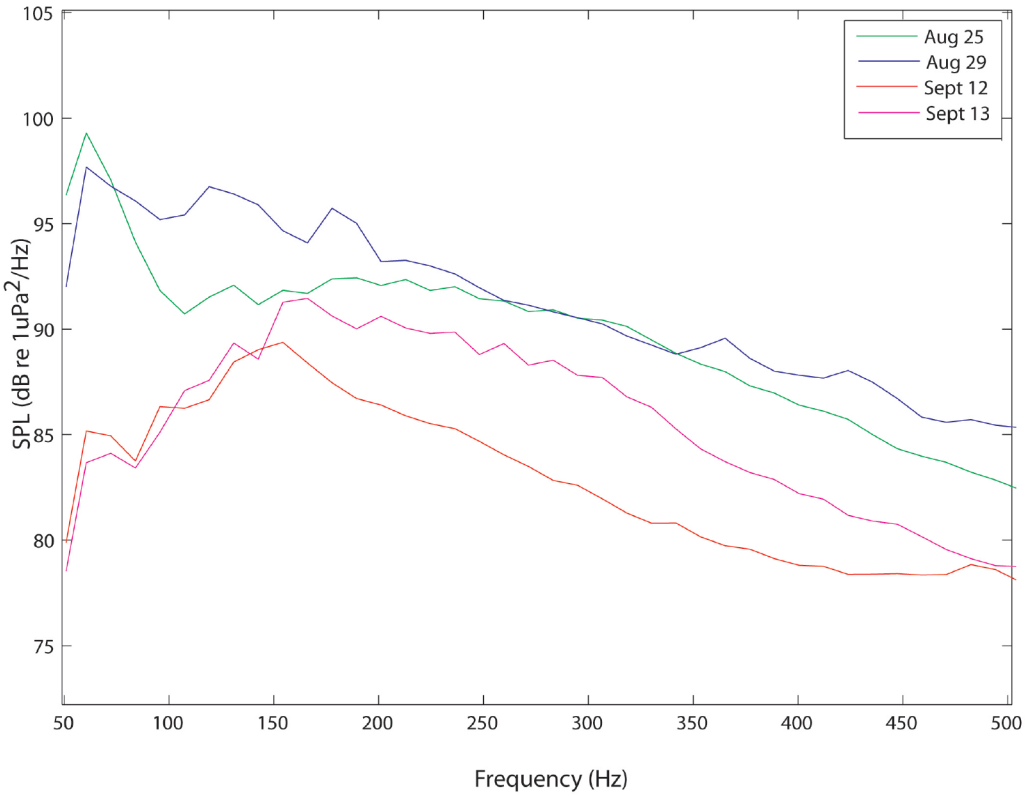

Figure 3. Spectrum level noise from 2 days before and 2 days after 9/11. Note the overall levels are lower, and the peak frequency of the noise has shifted to a higher frequency.