Forward Calculations:
Function and Procedure

Woods Hole Oceanographic Institution
Woods Hole, MA 02543

R. Levine, G. Lawson, and D. Chu

1 Introduction

1.1 Background

The following is a description of Dezhang Chu's MATLAB code for modelling backscatter from zooplankton of various shapes. The foundation of the code is the **zoo_bscat.m** function which executes the models using a vector of the animal lengths and a set of parameters used to specify the model type (**Section 2.3**).

The models can be used individually for a set of lengths, or through the forward calculation script for multiple taxa at multiple frequencies. While this manual will focus on the use of the script, be aware that all of the processes explained can be executed independently using their respective functions.

The forward calculation script included was originally designed to call the Load Files script and work with the output from WHOI Silhouette DIGITIZER, a MATLAB-based program for measuring marine organism lengths (See Little and Copley, 2003). The free software generates text files of the lengths and weights of the organisms as measured which can be imported into the Silhouette Template Excel workbook accompanying this manual, and used by the Load Files script to provide the lengths for individual taxa to the forward calculation script. More information on the technique of silhouette photography to create the images used within the the digitizing software can be found in Foote (2000), and in Davis and Wiebe (1985).

Additionally, many of the processes were designed specifically for the Multiple Opening/Closing Net and Environmental Sensing System (MOCNESS) and additional code specific to MOCNESS sampling is provided.

1.2 Accompanying Files

Forward Calculation Script The *Forward_Processing_Code.m* file performs the forward calculation models based on taxa-specific parameter settings and can execute the modelling for multiple nets containing multiple species, and at multiple frequencies. This script is the primary method of executing the forward calculations.

Load Files Scripts The *Load_Files_DIGI.m* file is used to read in the length vectors from a formatted Excel spreadsheet to the variable names expected by the forward calculation script. Alternatively, *Load_Files_taxa.m* can be used for data not in the WHOI Silhouette DIGITIZER output format. Both files contain an instruction sheet specifying the format expected. More on this in **Section 2.2.2**.

Code Library The *SubPrograms* directory contains all of the necessary functions that establish the parameters and execute the modelling of the acoustic backscatter.

Excel Data Template Along with the *.mat* files is an Excel spreadsheet used as the template for reading data into the forward calculations from the WHOI Silhouette DIGITIZER. Also provided is *FWD_Data_Template.xls*, a data template used by the *Load_Files_taxa.m* to assign the taxa variables.

Plotting Script Included is *Fwd_Preds_Plots.m* which uses the the observed median Sv for each net and the predicted Sv to create a series of plots showing consistency between the predicted and observed, and the percent contribution of backscatter from each group. The function and use of this file is explained in **Section 2.3**

Observed Sv Calculation The file *Fwd_Preds_Get_ObsSv.m* is used to calculate the backscatter for each net of a MOCNESS tow, using the concatenated acoustic data and the MOCNESS *.pro* file containing the time and sensor data for the tow. The function and use of this file is explained in **Section 2.4**

Example Data Included with the code is a silhouette data file (EN484M1_Silhouette_Data.xls), MOCNESS *.pro* file (M001.pro), and the concatenated acoustic data for the MOCNESS track (MOC1_Track_Concatenated.mat). Instructions on executing all of the MATLAB routines using the example data are provided in **Section 3**.

2 Forward Calculation

2.1 Tow and Net Information

Regardless of your sampling process, there are various names and labels within the code that are based on information such as cruise and net number (depending on sampling technique). As these MATLAB routines were originally designed to accompany samples collected using a MOCNESS, various defaults are provided which assume a similar method. The following will need to be changed:

Number of Nets A MOCNESS is designed to allow for multiple nets to be sampled during a single tow. To specify the number of nets, a value must be defined for the variable **'numnets'**. The default, assuming a standard MOCNESS tow is [**8**]. However, if you are only using a single net sample to create forward predictions, this value can just be set to [**1**]. In addition, if using a multiple net system, individual or non-sequential nets can be specified by identifying the columns in which the animal lengths are contained, such as by specifying [**1 5 7 2**] to process nets 1,2,5, and 7.

Frequencies: You can specify the frequencies to be used when executing the forward calculations. Two variables must be defined at the beginning of the forward calculation script to assign the frequencies. **'freqqs'** is a vector of the frequencies in kHz, such as [**43;120;200;420**]. In the initial **for** statement at the beginning of the loop, the index **'fq'** is defined to represent the same as the number of elements in your frequency vector.

Load Files The MATLAB routine used to load the animal lengths to their variable names, as expected by the forward calculation script, must be specified. Section **2.2** will explain how to format the length vectors.

Plot Figures Whether or not figures should be created must be specified at the beginning of the forward calculation code. Instructions and an explanation of the figures that can be produced and the necessary data are described in **Section 2.3**.

2.2 Organisms

2.2.1 Taxonomic Groups

The style of net sampling used will dictate how you have determined which animals are present and how organism lengths were calculated. The following taxonomic groups are represented within the MATLAB routines (*.m) used to calculate the forward predictions:

- Copepod
- Euphausiid
- Amphipod
- Fish
- Chaetognath
- Ostracod
- Pteropod
- Siphonophore
- Medusae
- Polychaete
- Salp
- Crustacean (larvae)
- Gymnosome Pteropod
- Mysid
- Larvacean
- Decapod

Also included within the routines are calculations for generic fluid spheres, such as eggs. In addition to these taxa, any animal that can be categorized to fit within the parameter options provided within the MATLAB routines can be included in the forward predictions (see **Section 2.3**). The default MATLAB routine requires the lengths of organisms in millimeters to be listed as separate variables in the form of a single column. The following sections describe various methods to create organism length vectors in the format used by the default routine.

2.2.2 Organism Length Files

Length measurements should be stored in column vectors. The **Load Files** script is used to take take an Excel sheet with multiple columns and assign that sheet to the corresponding variable name used in the forward calculation script. Within the script, the columns are reshaped into the vector L, used by the various parameters necessary to run **zoo_bscat**.

2.2.3 Importing Length Data From Excel Documents

For length information stored as a column within an Excel file, it is easiest to use the **xlsread** function. While the Load Files routine contained within this forward calculation instruction package is specifically designed to be used with the Excel template file provided, **xlsread** can be used with any style worksheet. The suggested use is:

$$\textit{Taxon Variable Name} = \text{xlsread}(\textit{`File Name'},\textit{`Sheet Name'},\textit{Range})$$

MATLAB will automatically fill in NaNs where xlsread finds cells containing strings. Additionally, when using xlsread to load an entire sheet, MATLAB will add NaNs to the end of columns as needed, in order to produce a rectangular matrix.

Specific instructions on the format of the data for the two Excel templates can be found on an instruction sheet in each file. Explanations of the formatting of the length vectors and specific directions on populating the spreadsheets for those using the data from the WHOI Silhouette DIGITIZER program are included.

2.3 Figures

2.3.1 Observed Acoustic Data

In order to create the plots, the mean or median acoustic backscatter for the sampling area must be calculated. The *Fwd_Preds_Get_ObsSv.m* script uses the MOCNESS track information provided in the *.pro* file to create the data from the raw acoustic data at each frequency. The script creates a polygon a fixed number of pixels in each direction from the track of the tow and bins a mean and median value for each net. The structure of the output is a matrix where each column corresponds to a frequency and each row a net, in sequential order, with *NaNs* for any missing values. The output is then saved into a *.mat* file loaded by the plotting script.

Fwd_Preds_Get_ObsSv.m uses the *m_lldist.m* function from the M_Map Toolbox for MATLAB. The toolbox is publicly available, however, the *m_11dist.m* function is provided in the example data folder for your convenience. More information on the M_Map toolbox is provided in **Section 5**.

2.3.2 Creating Figures

The figures can be created automatically at the end of the forward calculation by specifying **plots = '1'** at the beginning of the forward calculation script. **plots = '0'** indicates to turn plotting off. The script can be used independently, assuming all of the required variables from the forward calculation output and the observed data are either available in the workspace or within an active working directory.

The *Fwd_Preds_Plots.m* script uses the predicted Sv calculated by the models and the observed Sv (in the form of the median backscatter for each net of a MOCNESS tow). Figures are created to compare the observations and predictions for each taxonomic group modelled at each frequency specified by **freqqs**. Additional figures created also include the percent contribution of each taxon to the total backscatter, and comparisons of the total backscatter for each net.

Fwd_Preds_Plots.m uses the *fregres.m* function to calculate the slope, intercepts, and fit of the functional regression from the observed and predicted backscatter. The function was written by P.H. Wiebe adapted from Ricker (1973). The function is included in the folder along with the script for your convenience.

2.4 Setting Parameters

The following parameters are set within each subsection of the forward calculation script. The script simply uses **set_para.m** to assign the default parameter values for the **zoo_bscat.m** function, and then the parameters are changed based on the animal type, material properties, and any desired averaging for each taxa. **set_para.m** can be edited directly to adjust the default values for any of the following parameters.

2.4.1 Animal Type/Shape

para.simu.model This defines the animal body-type: 1 for fluid-like animals, 2 for a gas-filled sphere, and 3 for an elastic-shelled sphere.

para.simu.model_index For fluid-like animals, this defines the animal's shape: 1 for a cylinder or bent cylinder, and 3 for an ellipsoid. 2 is for two spheroidal interfaces (i.e., an umbrella-shaped animal such as a jellyfish)

para.shape.L_a This is the ratio of the animal's length to cylindrical radius, for cylinders or deformed cylinders. For ellipsoids, the animal in cross-section is described by two radii: one in a dorsal direction (a), and the other in a lateral direction (b). **para.shape.L_a** sets the ratio of the ellipsoid's length to dorsal radius (i.e., half its dorsal body width). For spheres, L/a is simply 2 (i.e., diameter/radius).

para.shape.L_b This only needs to be specified if the ellipsoid model is chosen, as described above. For a prolate spheroid, L_a = L_b.

2.4.2 Material Properties

para.phy.g1 The animal's density contrast.

para.phy.gL The animal's compressional sound speed contrast.

para.phy.cw The sound speed in water.

For elastic-shelled spheres, the material properties are parametrized as follows:

para.phy.g1 The density contrast of the outer shell relative to the surrounding water.

para.phy.g2 The density contrast of the animal interior of the shell relative to the surrounding water.

para.phy.hL The compressional sound speed contrast of the shell relative to the surrounding water.

para.phy.hT The transverse sound speed contrast of the shell relative to the surrounding water (necessary due to shear wave scattering).

para.phy.h2 The compressional sound speed contrast of the animal interior of the shell.

Average estimates for backscattering strength can be computed by averaging over some distribution of lengths, or angles of orientation:

para.simu.aveL_flag This sets whether or not to average over length. 1 for yes, 0 for no. If set to 0, the computed backscattering will only correspond directly to the exact lengths provided in the input.

para.simu.aveL_PDF Defines the type of length distribution, 1 for uniform or 2 for Gaussian.

para.shape.Lstd For Gaussian distribution, this defines the standard deviation. For uniform distribution, this parameter defines the half-range.

para.simu.nL Defines the number of discrete length increments to average over.

para.simu.aveA_flag Set whether to average over angle, 1 is yes, 0 is no.

para.shape.ang Define the mean angle of incidence in degrees. For the bent cylinder, 0 is broadside. For the ellipsoid, 0 is head-on and 90 is broadside (This is based on the assumption that the ellipsoid would be used for Calanoid copepods, which in the Gulf of Maine orient head-up). If you are not averaging over angle, this is the single angle of incidence.

para.simu.aveA_PDF Defines the type of angle distribution, 1 for uniform or 2 for Gaussian.

para.shape.dang For Gaussian distribution, this defines the standard deviation. For uniform distribution, this parameter defines the half-range.

para.simu.nA Defines the number of discrete angle intervals to average over.

In the case of the ellipsoid, the routines at present are not able to average over an angle, and so you have to set **para.simu.aveA_flag** to 0, and **para.shape.ang** to the single angle of incidence you are interested in. Since the ellipsoid can have different radii in its two cross-sectional planes, you also have to specify an azimuthal angle (ϕ) in radians. This is set using the following parameter:

para.shape.phi $= 0$ for sound striking the ellipsoid in dorsal aspect, or $= \pi/2$ for sound incident in lateral aspect.

2.4.3 Additional Parameters

para.simu.min_ni The minimum number of points for the numerical integration. This should be increased for longer animals and shorter wavelengths.

para.shape.L The maximum length of an animal in the length vector (i.e., max(L)). This parameter is used to calculate the maximum kL value which is then used to make sure that your entered **para.simu.min_ni** value is high enough.

para.simu.f0 and **para.simu.fe** set the frequency range in kHz, and in this application of the code must be set equal and must give the discrete frequency of interest. These values are used to calculate the wave number **para.simu.k**. Since this calculation is done within **set_para**, the frequency values **para.simu.f0** and **para.simu.fe** must likewise be set within the text of **set_para** itself, rather than modifying them after running **set_para**.

3 Executing the Forward Calculation

The following instructions are for executing the forward calculation using the example data provided. While specific to the example, the instructions are organized by each *.mat* file, covering what edits are necessary in each script, and can be used as reference for any dataset.

1. Edit Load_Files

The following variables need to be defined in Load_Files_DIGI.m:

CRUISE The cruise ID ('EN484').

MOC The MOCNESS tow number ('MOC1').

DIR The directory containing the Excel file of length data.

FILE_NAME The name of the Excel length file. In the case of the example data, this is the EN484M1_Silhouette_Data.xls file.

If you are using *Load_Files_taxa.m* for your own work, the same variables must be defined. Similar instructions can be found in the header of the .m files.

2. Generate Observed Acoustic Data

If you would like to create the accompanying figures for the forward calculation, the acoustic data must be generated and available before running the forward calculation script. The following variables must be defined in the *Fwd_Preds_Get_ObsSv.m* script:

nets This must be set the amount of nets available in the *.pro* file associated with the MOCNESS tow. In the case of the example data, **nets = 7**.

dirna_acc The directory containing the concatenated acoustic data for the tow.

filena_acc The name of the concatenated acoustic file.

MOCfull You must specify the path and name of the *.pro* file for the MOCNESS tow.

M_map Toolbox If not already in your path or added to your default toolbox, you must specify the location of either the M_map Toolbox or the *m_11dist.m* function.

Once complete, the last thing to edit is the name of the *.mat* file that will be created at the end containing all of the output. This is specified in the **save** statement before the last line of the script. The .mat file will be directly loaded if plots are turned on the in the forward calculation script.

3. Plot Script

The *Fwd_Preds_Plots.m* script has a few changes to be made if it is going to be run from within the forward calculation script.

Cruise and Tow ID The cruise ID and tow number need to be set to be consistent with the forward calculation script (this will be discussed in the next step). The easiest method is to use *CTRL + f* to find and replace all of the cruise ID and tow name prefixes with the current data you are using. For example, in the script accompanying the example data, you will notice that the beginning of all of the backscattering variables start with *'EN484M1'*. You may wish to change this by simply executing a *replace all* from within the find command window for your own cruise name.

Mean or Median You can specify whether you would like to use the mean backscatter or median backscatter as the observed data when creating the plots. This is done by assigning either **ac_mean** or **ac_med** to the *'CRUISE'_acc* variable at the top of the script.

4. The Forward Calculation

The final step is to edit and then run the forward calculation script, *Forward_Processing_Code.m*. As discussed in **Section 2.1**, there are a few variables that need to be defined, as well as the path of the models to be set:

SubProgram Directory The *SubProgram* directory and its subdirectory *lib* must be added to the path of the forward code, and can be done so following an *addpath* command.

Load Files The *Load_Files_DIGI.m* or *Load_Files_taxa.m* must be specified. This is the first thing called within the script to assign all of the length vectors to their variables. The example data is stored in the Excel format used for the silhouette output, and therefore *Load_Files_DIGI.m* is used.

plots plots can be set to **'on'** or **'0'**. If **'1'**, after all of the models for each frequency is complete, the script will call *Fwd_Preds_Plots.m*.

freqqs The frequencies must be specified as a vector. The same assignment for **freqqs** will be used to create the figures as well.

numnets The number of nets, relating to the columns of the silhouette data, need to be specified. Here, the example data contains 7 nets, so **'numnets = 7'**

4 References and Related Readings

Anderson, V. C. 1950. Sound scattering from a fluid sphere. The Journal of the Acoustical Society of America, 22: 426-431.

Chu, D., Wiebe, P., and Copley, N. 2000. Inference of material properties of zooplankton from acoustic and resistivity measurements. ICES Journal of Marine Science,57: 1128-1142.

Chu, D., Wiebe, P. H., Copley, N. J., Lawson, G. L., and Puvanendran, V. 2003. Material properties of North Atlantic cod eggs and early-stage larvae and their influence on acoustic scattering. ICES Journal of Marine Science, 60: 508-515.

Davis, C. S., and Wiebe, P. H. 1985. Macrozooplankton biomass in a warmcore Gulf Stream ring: Time series changes in size structure, taxonomic composition, and vertical distribution. Journal of Geophysical Research, 90: 8871-8884.

Foote, K. G. 2000. Optical Methods. *In* Zooplankton Methodology Manual,pp. 259-295. Ed. by R. P. Harris, P. H. Wiebe, J. Lenz, H. R. Skjoldal, and M. Huntley. Academic Press, San Diego.

Kristensen, A., and Dalen, J. 1986. Acoustic estimation of size Distribution and abundance of zooplankton. The Journal of the Acoustical Society of America, 80: 601-611.

Little, W. S., and Copley, N. J. 2003. WHOI Silhouette DIGITIZER Version 1.0 Users Guide. Woods Hole Oceanographic Institution Technical Report # 2003-05, 63 pp.

Medwin, H., and Clay, C. S. 1998. Fundamentals of Acoustical Oceanography. Academic Press, New York. 712 pp.

Ricker, W. E. 1973. Linear regressions in fishery research. Journal of the Fisheries Board of Canada, 30: 409-434.

Stanton, T. K., and Chu, D. 2000. Review and recommendations for the modelling of acoustic scattering by fluid-like elongated zooplankton: euphausiids and copepods. ICES Journal of Marine Science, 57: 793-807.

Stanton, T. K., Chu, D., and Wiebe, P. H. 1998. Sound scattering by several zooplankton groups. II. Scattering models. The Journal of the Acoustical Society of America, 103: 236.

Stanton, T. K., Wiebe, P. H., Chu, D., Benfield, M. C., Scanlon, L., Martin, L., and Eastwood, R. L. 1994. On acoustic estimates of zooplankton biomass. ICES Journal of Marine Science, 51: 505-512.

US SO GLOBEC, 2002. Report of RVIB Nathaniel B. Palmer Cruise NBP02-04 to the Western Antarctic Peninsula 31 July to 18 September 2002. United States Southern Ocean Global Ocean Ecosystems Dynamics Program Report Number 8. Old Dominion University, Norfolk, VA.

5 Additional Resources

M_Map

M_Map is a mapping package for MATLAB created, maintained, and distributed by Rich Pawlowicz. The mapping tools are available for MATLAB v5 and later. All of the contents of the toolbox are publicly available at http://www.eos.ubc.ca/ rich/map.html.

Organism Parameters			Default Parameter values			
Variable Name	*Taxon*	*Scattering Model*	*Orientation*	*Density Contrast (g)*	*Sound Speed Contrast (h)*	*Length to Width Ratio*
'copslen'	Small copepod	DWBA Bent Cylinder Model[1]	N(0,30)	1.02[6]	1.058[13]	2.5497
'copllen'	Large copepod	DWBA Bent Cylinder Model[1]	N(0,30)	1.02[6]	1.058[13]	2.5497
'eudelen'	Large Euphausiid (acoustic length)	DWBA Bent Cylinder Model[1]	N(0,27)	1.039[7]	1.038[7]	8.51539468
'euphslen'	Euphausiid	DWBA Bent Cylinder Model[1]	N(0,27)	1.0157[8]	1.0189[8]	9.2
'amphlen'	Amphipod	DWBA Bent Cylinder Model[1]	N(0,30)	1.058[9]	1.058[9]	3.0021
'fishlen'	Fish	DWBA Bent Cylinder Model[1]	N(0,30)	1.03[6]	1.03[6]	4
'chaelen'	Chaetognath	DWBA Bent Cylinder Model[1]	N(0,30)	1.03[6]	1.03[6]	17.151
'ostrlen'	Ostracod	DWBA Bent Cylinder Model[1]	N(0,30)	1.03[6]	1.03[6]	2.5497
'limalen'	Pteropod	Spherical Elastic Shell Model[2]	-	1.732[15]	1.732[15]	-
'sphrlen'	Sphere (i.e., egg)	Spherical Elastic Shell Model[2]	-	0.979[10]	1.017[10]	-
'snectlen'	Siphonophore - Nectophore	Fluid Sphere Model (equivalent volume sphere)[4]	-	1.02[6]	1.02[6]	-
'spneulen'	Siphonophore - Pneumatophore	Gas Sphere Model (depth dependent density)[5]		$g=g_{surface}(1+.01Depth(m))$[12]	0.22[14]	-
'medlen'	Medusae	DWBA-based Model of Two Spheroidal Interfaces[3]	N(0,20)	1.02[3]	1.02[3]	-
'sbractlen'	Siphonophore - Bract	Fluid Sphere Model (equivalent volume sphere)[4]	-	1.02[6]	1.02[6]	-
'polylen'	Polychaete	DWBA Bent Cylinder Model[1]	N(0,30)	1.03[6]	1.03[6]	17.151
'salplen'	Salp	DWBA Bent Cylinder Model[1]	N(0,30)	1.0041	1.0041	4
'crustlen'	Larval Crustacean	DWBA Bent Cylinder Model[1]	N(0,30)	1.058[9]	1.058[9]	2.5497
'ptergymnlen'	Gymnosome Pteropod (i.e., Clione)	DWBA Bent Cylinder Model[1]	N(0,30)	1.03[6]	1.03[6]	1.83
'radioforamlen'	Radiolarian and Foram	Spherical Elastic Shell Model[2]	-	2.147[11]	3.979[11]	-
'mysidlen'	Mysid	DWBA Bent Cylinder Model[1]	N(20,20)	1.002	1.004	5.3576
'lvcnlen'	Larvaceans	DWBA Bent Cylinder Model[1]	N(0,30)	1.03[6]	1.03[6]	1.83
'decalen'	Decapod	DWBA Bent Cylinder Model[1]	N(0,30)	1.039	1.038	10.32

Table 1: Organism Names and Defaults - This table is a modified version of the one found in Lawson et al. (2004), Deep-Sea Research II, 51: 2041-2072.

[1] Equation (5) of Stanton et al. (1998), and see Stanton and Chu (2000) and references therein.

[2] Equation given in Stanton et al. (1994), p. 507.

[3] D. Chu. Unpublished Data

[4] D. Chu. and A. Lavery, Pers. Comm.; Fluid-filled sphere model is derived from Anderson (1950)

[5] Derived from Anderson (1950)

[6] D. Chu. Pers. Comm.

[7] From Lucio Calise's dissertation (similar to Kristensen and Dalen, 1986)

[8] D. Chu. Unpublished Measurements; preliminary descriptions of results are found in U.S. SO GLOBEC (2002). Measurements were performed only on animals larger than 20 mm. For animals smaller than this length (e.g., the small euphausiid category), the g and h predicted by the regression equations for a 20 mm animal were assumed to apply.

[9] Inferred for shrimp based on model-fits to direct observations by Stanton et al. (1994), and very comparable to values measured by Chu et al. (2000) for the decapod shrimp Palaemonetes vulgaris. Assumed here to also apply to certain other crustacean and crustacean-like taxa.

[10] Measured by Chu et al. (2003)

[11] Based on the density and sound speed of fused silica (2.2 g m-3 and 5968 m s-1, respectively), and assuming the speed of sound in seawater is 1500 m s-1, and the density of seawater is 1.025 g m-3.

[12] Where $g_{surface}$=0.0012, From A. Lavery. Pers. Comm.; based on pressure-related increases in density and thereby g in depth (Medwin and Clay, 1998). g surface is the density contrast for carbon monoxide at the surface (1 atmospheric pressure).

[13] Measured by Chu et al. (2000) for Gulf of Maine calanoid copepods

[14] Sound speed contrast for carbon monoxide at surface pressure of 1 atmosphere

[15] Inferred based on model-fits to empirical observations, Stanton et al. (1994)