Chapter 2: Methodology

1. Theory

The theoretical concept underlying the PAWSA process is the proven Delphi method of converting the opinions of local subject matter experts into quantified results. This method is used so that those quantified results can be compared internally (i.e., the results for one risk factor can be compared to those for other risk factors and the results from one stage (e.g., *Book 1*) can be compared to the results from other stages (e.g., *Book 3*) during the workshop) and externally (i.e., the results from one waterway can be compared to the results from other waterways). The strength of the PAWSA process derives from several sources:

(1) the sponsor carefully selects the participants because they are knowledgeable with respect to a particular maritime interest and so that all important interests are represented within the group;

(2) before converting their opinions into numbers, the participants thoroughly discuss the issues being judged;

(3) the same 1 to 9 scale is used repeatedly throughout the process; and

(4) all quantified inputs are weighted by the relative expertise of each participant team with respect to each risk category in the Waterway Risk Model.

Proof that the PAWSA process works (i.e., produces valid results) comes from the internal consistency checks that are built into the results spreadsheets within the Excel™ workbook (PAWSA software) used to capture and analyze the participants' quantified inputs. Those consistency checks have repeatedly shown that workshop participants develop strong consensus about the levels of risk in the waterway and the effectiveness of various risk mitigation strategies. This consensus emerges in spite of the fact that the participants typically represent widely different interests within the overall maritime community and in spite of the fact that the 1 to 9 measurement scale used is correlated only loosely with qualitative descriptors for each value on that scale. (See Chapters 6 and 7 for elaboration.)

The rest of this chapter specifically describes *Books 1 – 4*. Understanding how each book is used, the methodology behind each book, and how the PAWSA software relates to each book are critical to understanding the overall PAWSA process.

2. Book 1: Establishing Baseline Risk Levels

To determine a risk level value for every factor in the Waterway Risk Model, *Book 1: Baseline Risk Levels* uses the same four qualitative descriptors for each risk factor as were used in *Book 2*. In theory those qualitative descriptors are written in absolute terms; that is, the risk level values that are produced by *Book 3* do NOT take into account any actions already implemented to reduce risk in the waterway. In practice, PAWSA participants sometimes have difficulty thinking in such absolute terms and the effects of existing mitigations tend to creep into the discussion and evaluation of this workshop stage.

Key to achieving strong consensus in the *Book 3* results is the discussion period that immediately precedes filling out this quantitative evaluation. During that discussion the various perspectives concerning each risk factor are voiced and, sometimes, debated. Often participants refer to read-ahead

material provided by the sponsor (or readily available to them via other means), especially for risk factors amenable to measurement and/or quantification (e.g., volume of traffic, wind conditions, cargo volumes). Once the discussions have run their course, participants simply check the box next to the qualitative descriptor for a particular risk factor that best matches conditions in the waterway being evaluated.

If a team checks the first box (describing the best case), then a 1 is entered into the *Bk 3 Input* spreadsheet, obviously corresponding to a value of 1.0 for that input. If a team checks the second box, then a 2 is entered into the spreadsheet and the computer algorithm assigns the "B" value from the aggregate risk measuring scale for that factor to that input. In like manner, a check in the third box is entered as a 3 and assigned the "C" value; a check in the fourth box (describing the worst case) is entered as a 4 and assigned a value of 9.0.

Building on the same three team Wind Conditions example from previous sections, hypothesize the following *Book 3* inputs:

	Team 1	Team 2	Team 3
Box Checked	Third	Second	Third
Spreadsheet Entry	3	2	3
Risk Value	C	B	C
Value Assigned	5.50	2.59	5.50

The inputs for each team for each factor are multiplied by their team expertise scores and then added together to produce the baseline risk value for that factor. Continuing our example:

	Team 1	Team 2	Team 3	Sum
Value Assigned	5.50	2.59	5.50	
Expertise Score	.43	.14	.43	
Product	2.36	.36	2.36	5.08

Thus, for this example, the baseline risk value for the Wind Conditions factor is 5.1. (Note: All results are displayed rounded to one decimal place because the qualitative descriptors that underlie these quantitative results are not precise enough for greater numerical precision.) The results from *Book 3* for each risk factor in the Waterway Risk Model become the baseline from which the effectiveness of existing mitigation strategies are evaluated in *Book 4*. Those baseline numbers are marked on the *Book 4* assessment forms using a highlighter pen.

3. Book 2: Team Expertise Cross-Assessment

There is no expectation that every participant invited to a PAWSA workshop will be equally knowledgeable with respect to all of the risk factors included in the Waterway Risk Model. The 24 risk factors in the model were developed to provide the foundation for discussions that could include the very broad range of maritime safety issues found throughout the United States and internationally.

Because the PAWSA participants are expected to have varying expertise with respect to the risk categories in the Waterway Risk Model, *Book 2: Team Expertise Cross-Assessment* is used early in the session (towards the end of Day 1) to weigh the relative strengths of each team with respect to the six risk categories. After being presented with the concepts underlying the model, each participant team is asked to discuss (among themselves) how their background and experience aligns with the model. They then verbally present their conclusions to the larger group. This presentation gives all

teams a sense of where everyone thinks they are strong – or perhaps not so strong. After all teams have spoken, each team evaluates whether they think they are in the top, middle, or lower third of all teams present in knowledge about the six risk category areas. Throughout the workshop, these expertise evaluations are used to produce results for all other Books.

The teams, in doing the expertise evaluation, conceptually are dividing up six expertise pies (risk categories) into different sized slices, with the relative size of each slice from each pie equaling the expertise of each team relative to the other teams for that risk category. An example for the Navigational Conditions Risk Category:

Team 1: Circles a 1 indicating they are in the Top 1/3 of all teams present

Team 2: Circles a 3 indicating they are in the Lower 1/3 of all teams present

Team 3: Circles a 1 indicating they, also, are in the Top 1/3 of all teams present

These responses are entered into the data input cells in the *Bk 2 Input* spreadsheet in the *All Books (waterway name)* workbook. Those inputs are inverted, i.e., all inputs are subtracted from 4 so that a 1 becomes a 3 and a 3 becomes a 1. This is done so that the Top 1/3 teams get the biggest slice of the pie. Those inverted scores are added up (showing that, in this case, the total pie size = 3 + 1 + 3 = 7). Then each team's slice is computed by dividing their inverted score by the total pie size. For our example:

Team 1: 3/7 = .429 (≈ 43% of the Navigational Conditions expertise pie)

Team 2: 1/7 = .143 (≈ 14% of the Navigational Conditions expertise pie)

Team 3: 3/7 = .429 (≈ 43% of the Navigational Conditions expertise pie)

Obviously, but mathematically very important, adding all of the slices together equals 100% of each expertise pie. These computations are done independently for each of the six risk categories (expertise pies). Each team's relative expertise in each category (size of their slice) is multiplied by their inputs for the four risk factors <u>in that category</u> during all of the other quantitative evaluations (*Books 3* and *4*). When this multiplication is done, the products that result are the weighted inputs for that team for that book. Because the sum of the expertise for each category equals 100%, the sum of the weighted inputs equals the risk level.

4. <u>Book 3: Mitigation Effectiveness</u>

Again, the key to good consistency in results from the *Book 3: Mitigation Effectiveness* stage is the discussion that immediately precedes filling out the quantitative evaluations. Those discussions should focus on three issues: (1) the specifics of what has been done to reduce the risk associated with a particular factor; (2) the effectiveness of those mitigation actions; and (3) whether existing mitigations are well balanced with the baseline risk value.

Once the discussions are complete, the participants are asked to do two things: (1) circle a number on the 1 to 9 scale that shows the effectiveness of existing mitigations in reducing risk below the absolute levels determined via *Book 1* and (2) circle Yes (or No) depending on whether they think existing mitigations adequately balance the risks for each factor (or not).

The vast majority of the time, participants will circle a number on the 1 to 9 scale to the left of (smaller than) the highlighter mark denoting the *Book 1* result. However, if they conclude that actions taken previously are having no effect on reducing the baseline risk, they will circle the *Book 1* result

mark. Though unusual, participants might state (and then evaluate) that existing mitigations actually increase the risk for some factor(s). For example, while discussing the Dimensions risk factor, participants cite as an existing risk mitigation strategy that a range light has been established to help waterway users keep from running aground in a narrow channel, but state that the range is out of alignment with the channel, thereby increasing the risk of groundings. They then could evaluate the effect of that mitigation by circling a higher number (i.e., to the right) of the *Book 1* result mark.

The numbers that are circled by the participants are entered exactly as indicated into the *Bk 3 Scores* spreadsheet with two exceptions: (1) if the participants circle the space between two whole numbers, the entry is invalid and the team is required to reassess providing a whole number entry; and (2) if the participants circle the *Book 1* result mark, a lower case "e" is entered and the computer algorithms convert that entry into the *Book 1* results value.

As with *Books 1 and 4*, *Book 3* numerical entries are multiplied by the *Book 2* expertise scores and then those products are added together to produce the present risk level, which takes into account the effectiveness of existing mitigations.

Continuing our example from previous sections:

	Team 1	Team 2	Team 3	Sum
Number Circled	3	Highlighter mark	4	
Spreadsheet Entry	3	e	4	
Value Assigned	3	5.08	4	
Expertise Score	.43	.14	.43	
Product	1.29	.71	1.72	3.72

Rounding this result to one decimal place, we see that the effectiveness of existing mitigations in reducing Wind Conditions risk is judged to be: 5.1 – 3.7 = 1.4 points.

As the final step in *Book 3*, participants make a subjective evaluation, based on the preceding discussions, as to whether they think risks are adequately balanced with existing mitigations for each factor. They do this by circling Yes (they are balanced) or No (they are not balanced) on the line in *Book 3* for each factor. Those Yes / No answers are coded into the *Bk 3 Y-N* spreadsheet as lower case "y" or "n". If 2/3 or more of the participant team expertise indicates Yes, then that risk factor is dropped from further discussion / evaluation in *Book 4: Additional Interventions*. This condition is denoted by a green *Balanced* on the *Book 3* results display spreadsheet (*Bk 3 Disp*). If 2/3 or more of the participant team expertise indicates No, then that risk factor should definitely be discussed / evaluated in *Book 4*. That condition is denoted by a red *NO* on the *Book 3* results display. If there is less than 2/3 consensus about the efficacy of existing mitigations then a yellow *Maybe* is displayed. Those "Maybe" risk factors should also be discussed / evaluated in *Book 4*. Finally, if the present risk level is evaluated as being HIGHER than the risk level from *Book 1* or, when appropriate, is higher than the risk level determined during a previous PAWSA held for the same waterway (see *Appendix A: Converting Risk Model Scores*), then a red *RISING* is shown on the *Book 3* results display.

5. Book 4: Additional Interventions

In the final quantitative evaluation stage of the PAWSA process, discussion is focused on those risk factors where the present risk level is not *Balanced*. For each risk factor displaying a *NO*, *RISING*, or *Maybe* flag, the *Book 3* results are marked using a highlighter on blank copies of the *Book 4: Additional Interventions* evaluation forms. This serves as a starting point for evaluating the possible effectiveness of new mitigation strategies. For each risk factor so marked, the workshop participants

are asked to offer ideas about what should be done to reduce the present risk level. Again, the quality of the discussion directly affects consistency of results obtained.

Analysis of risk mitigation ideas offered to date showed that those ideas usually fall into nine major implementation categories. Those categories are:

- Coordination / Planning
- Voluntary Training
- Rules & Procedures
- Enforcement
- Navigation / Hydrologic Information
- Radio Communications
- Active Traffic Management
- Waterway Changes
- Other Actions

Those categories are fully described in Chapter 6, are provided in each participant's folder under *Appendix H: Glossary of Terms* (see Chapter 5), and also are defined on the first page of *Book 4*.

After the participants have presented / discussed their risk mitigation ideas, they are asked to write short phrases (3 to 5 words each) describing the ideas they think have merit. Those short phrases are written on the lines next to the categories into which the ideas best fit. For example, if the risk factor being discussed is Wind Conditions and the idea being considered is "Install wind sensor at Long Point", then the participants would write those words on the line next to the Nav / Hydro Info intervention category under that risk factor. After recording an idea, the participants indicate what risk level would result from implementing that idea. This is done by circling a number to the left of (lower than) the *Book 3* risk level mark on the 1 to 9 scale next to the implementation category where the idea was written. As in *Book 3*, only whole numbers are used; therefore, if the participants circle the space <u>between</u> two whole numbers, the entry is invalid and the team is required to reassess providing a whole number entry. The closer that circle is to 1, the more effective the participant team feels the idea to be. Those evaluations are again multiplied by the team's expertise scores and then

those products are added together to get the possible risk level resulting from implementing the ideas written down for a particular category.

Again using our Wind Conditions example:

	Team 1	Team 2	Team 3	Sum
Number Circled	2	3	2	
Spreadsheet Entry	2	3	2	
Expertise Score	.43	.14	.43	
Product	.86	.42	.86	2.14

The algorithms for the *Book 4* display spreadsheet (*Bk 4 Disp*) determine which implementation category most teams have chosen and then how much risk improvement would result from the ideas

written down for that category. Those *Book 4* display algorithms also determine which implementation category was judged to be most effective (i.e., had the biggest delta between the *Book 4* and *Book 3* results). A yellow *Caution* flag is displayed if the most chosen implementation category is NOT the same as the most effective category AND either fewer than 50% of the teams chose the most chosen category OR more than 50% of the teams chose the most effective category. The presence of the yellow *Caution* flag for any risk factor indicates the possibility that there is more than one "best" mitigation measure to use to achieve further risk reduction for that factor.