Journal of Archaeological Science 31 (2004) 1533–1552

Journal of Archaeological SCIENCE

http://www.elsevier.com/locate/jas

The identification of shipwreck sites: a Bayesian approach

John M. O'Shea*

Museum of Anthropology, University of Michigan, Ann Arbor, MI 48109-1079, USA

Received 4 February 2004; received in revised form 18 March 2004

Abstract

Archaeology of the historic era offers the promise of linking people and events to archaeological materials. Yet such identifications have an 'all or nothing' quality and, once made, often take on a reality all their own that does not convey the underlying degree of certainty or data quality on which the identification was based. This is a common problem in archaeology, where both the data, as well as the conclusions, are inherently probabilistic. This paper focuses on the analysis of archaeological shipwrecks, and develops a system to assess and quantify the level of confidence that can be attached to the association of a historical vessel with a scattered wreck site. The linking of a historic vessel with the wreck site is treated as a probabilistic process, and Bayesian methods are employed to estimate the confidence with which a particular vessel is linked to a wreck site and to update this confidence as new evidence is introduced. The approach is demonstrated on a series of wreck sites from the Great Lakes region.

Keywords: Shipwrecks; Bayesian analysis; Great Lakes

1. Introduction

One of the benefits for archaeologists working on sites from the historic era is the potential of linking the discovered materials with the rich data that can be obtained from written sources. This is particularly true for the archaeological study of shipwrecks, where a vessel identification may provide immediate links to construction plans of the vessel, the names of the captain and crew, the manifest of cargo, and the often sad and harrowing tale of the vessel's loss. Historical records can also play a critical role in locating and identifying the wreck site, by providing descriptions of the vessel and the place, time, and conditions of loss.

Given the great value that accrues from linking material remains to historic accounts, archaeologists of the historic era work diligently to make the connection and, in the context of shipwreck archaeology, to identify the name of the vessel. Yet in this process, vessel identification tends to become an 'all or nothing' proposition. That is to say, the material remains are either unidentified or they *are* the named vessel. In this process, the inherently probabilistic nature of the identification is obscured. The true level of confidence or uncertainty with which the identification was made is lost, as is the possibility that other vessels might also make a plausible match with the archaeological remains. While every archaeologist and historian knows that identifications vary in terms of the strength of evidence and reliability, once the linkage is made, the wreck becomes that boat.

Problems of this kind are common in shipwreck archaeology. A recent example is found in the much-touted discovery of the steamer *Chicora*, a vessel that was tragically lost on Lake Michigan with great loss of life in 1895. Sophisticated models of the conditions during the storm and lake circulation, coupled with intensive remote sensing survey succeeded in discovering an intact vessel [27]. Unfortunately, subsequent research demonstrated that the vessel was not the *Chicora*, and is more likely the steamer *H. C. Akeley*, which was lost in the same area in November of 1883. The error in

* Tel.: +1 734 764 0485; fax: +1 734 763 7783.
E-mail address: joshea@umich.edu

0305-4403/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jas.2004.03.015

identification only became apparent when the deck layout of the wreck was found to be incompatible with the construction plans of the *Chicora* [28].

The potential for erroneous identification is greatly increased in the study of scattered wrecks [21,23], where the vessel is no longer intact and its characteristics must be inferred from the analysis of the remaining pieces of the vessel. Likewise, the potential for correcting erroneous identifications is much more limited, since one can rarely hope to discover an associated name on the vessel or other evidence that would conclusively undermine an existing identification. The particular problems posed by the study of scattered wreck sites are similar to those faced by any prehistoric archaeologist, where the characterization of the past community or society must itself be built up through an inherently probabilistic process of archaeological inference using incomplete and often imprecise data. While a wreck site can yield much important historical and archaeological evidence even when its name is unknown, the information potential is vastly increased when the physical evidence can be linked to contemporary historical documents. It is also true that even when a vessel cannot be identified with any scientific certainty, links will still be claimed to some known vessel by divers and local historians. The mere repetition of such claims is often sufficient to create a reality all its own.

In an effort to address these issues, this paper describes a system, *Bayes Boats 1*, that supports a probabilistic model of the inference process tied to the association of a historical vessel with a scattered wreck site. The system employs a Bayesian approach to quantify the relative confidence that can be placed in a given wreck identification and permits a numerical comparison of the relative likelihood that the wreck might be attributed to any one of a series of other plausible historical vessels. Of particular importance for the archaeological study of scattered wreck sites, these relative probabilities are updated in light of new evidence. In the sections that follow, a brief introduction to Bayesian methods is presented, followed by a description of the Mark I version of the assessment engine. The potentials of the system are explored using two archaeological cases from the Great Lakes region of North America.

2. Bayes' Theorem and archaeology

Thomas Bayes was an English clergyman and mathematician. His principal contribution to modern statistics was an essay entitled 'Essays Towards Solving a Problem in the Doctrine of Chances' which was published posthumously in the *Philosophical Transactions of the Royal Society of London* in 1763. From this work comes what is now known as Bayes' Theorem.

$$P(\mathrm{H}|\mathrm{E}) = P(\mathrm{H})^* P(\mathrm{E}|\mathrm{H})/P(\mathrm{E})$$

This formula states that the probability of a hypothesis given a specific piece of evidence ($P(\mathrm{H}|\mathrm{E})$) is equal to the simple probability of the hypothesis ($P(\mathrm{H})$) times the probability of the piece of evidence given the hypothesis ($P(\mathrm{E}|\mathrm{H})$) divided by the simple probability of the evidence ($P(\mathrm{E})$). For the calculations used in the present application, a variant of this formula is used.

$$P(\mathrm{H}|\mathrm{E}) = P(\mathrm{H})^* P(\mathrm{E}|\mathrm{H})/[P(\mathrm{H})^* P(\mathrm{E}|\mathrm{H}) + P(\sim\mathrm{H})^* P(\mathrm{E}|\sim\mathrm{H})]$$

This states that the probability of a hypothesis given a specific piece of evidence ($P(\mathrm{H}|\mathrm{E})$) is equal to the simple probability of the hypothesis ($P(\mathrm{H})$) times the probability of the piece of evidence given the hypothesis ($P(\mathrm{E}|\mathrm{H})$) divided by these same values [$P(\mathrm{H})^*P(\mathrm{E}|\mathrm{H})$ plus the probability *not* the hypothesis ($P(\sim\mathrm{H})$) times the probability of the evidence given *not* the hypothesis ($P(\mathrm{E}|\sim\mathrm{H})$)].

In its simplest form, and stripped of all metaphysics, Bayes' Theorem is simply a mathematical formula that is used to calculate and update conditional probabilities [13]. It provides a means to quantify how belief or confidence in a hypothesis changes with the introduction of new information. It starts with an initial level of belief, stated as a probability, which is then modified as new evidence is introduced. This starting point is referred to as the *prior* probability, while the updated value is conventionally termed the *posterior* probability. The posterior probability may in turn serve as a prior probability as new evidence is accumulated [17]. Such an evaluation of probability is not restricted to a single hypothesis. For example, we may want to decide among three possible explanations for some phenomenon. In this case, the prior probability for each explanation would be specified and these values would all be updated as new information is acquired. Such an approach seems ideally suited to archaeological research and would even seem to offer a mechanism for quantifying the method of multiple working hypotheses that is so often advocated for archaeological research [25,30].

Since its initial formulation, there have been practical and metaphysical objections to the Bayesian approach. Among these are such fundamental issues as the relationship between deductive versus inductive reasoning and the objective versus subjective nature of science, as well as the question of whether Bayesian probabilities are actually probabilities at all. For the purposes of the present archaeological application, however, these great issues are not of particular importance. Instead, this study adopts what Joyce [13] terms a 'minimal Bayesian' approach as a basis for identifying hypotheses (historic vessels) that are best supported by the data (characteristics of a particular archaeological wreck site). Yet, if the high-level philosophical issues are not of particular concern, some of the more practical issues are important.

Perhaps the single most common criticism of Bayesian methods is the need to establish prior probabilities. This is viewed by some as introducing an unacceptably subjective element into the analysis, which undermines the credibility of any conclusions [16]. Yet, it is hard to imagine a situation where the investigator starting a research project does not have some intuitive idea, based on previous research, reading, or experience of the relative strength or plausibility of the various potential explanations [7]. Archaeologists constantly make assessments of the most likely explanation or cause for the materials observed. While these judgments are rarely quantified, they are made nonetheless. To apply a Bayesian approach it is necessary to estimate how likely we believe each explanation to be. Alternatively, we can limit our prior knowledge to a simple identification of plausible versus implausible explanations, and consider each plausible explanation as having an equal prior probability. Of course, the mere act of deciding which potential explanations warrant consideration is itself a statement of prior probability among the universe of potential explanations.

In practical terms, while there is a great deal of discussion regarding the specification of prior probabilities, the initial values have relatively little impact on the posterior values once a quantity of evidence has been accumulated [12]. Aside from philosophical and metaphysical issues, the primary value of establishing realistic prior probabilities is analytical efficiency.

Bayesian analysis has seen scattered application in archaeology. By far the most common use in archaeology has been in the area of interpreting radiocarbon dates and chronology (cf. Refs. [4–6]). Other applications have been in areas such as remote sensing [2] and artifact classification (cf. Refs. [8,10,18]).

In the present application, Bayes' Theorem is used to quantify the relative confidence with which a particular historical vessel can be linked to an observed archaeological wreck site. In conventional Bayesian terminology, each of the plausible historic vessels constitutes a hypothesis, whose prior probability is updated conditional to the evaluation of new evidence, which derives from archaeological observations at the wreck site. In the two cases that will be presented, the analysis begins with a set of historic vessels that are thought to have been lost in the region. Within this set of plausible identifications, each historic vessel is assigned an equal prior probability of being the source of the wreck site. The analysis then updates the probabilities for each historic vessel in light of the observable characteristics of the archaeological remains across five sets of observations. The final posterior probabilities reflect the relative confidence that can be placed in each historic ship as the potential source of the wreck.

3. Bayes Boats 1

The starting point for the development of the model, termed here *Bayes Boats 1*, was a preliminary cataloging of the range of historical information that typically is available for nineteenth and early twentieth century wooden vessels on the Great Lakes, and an assessment of the types of archaeological evidence that are preserved in scattered wreck sites. Of central importance was the identification of those aspects where the archaeological evidence can be compared with the historical descriptions. The model excluded vessels constructed of iron and steel, as well as primarily ocean going vessels, although each could easily be accommodated under a similar modeling process.

On the historical side of the ledger, most commercial vessels operating on the Great Lakes during the nineteenth and twentieth centuries underwent registration, at the time of their commissioning, and throughout their use life, often culminating in a final document reflecting the loss or decommissioning of the vessel. The registry documents provided information on the size, capacity, and type of ship along with details of ownership, homeport and, in the case of loss, the cause of loss. Registry information is supplemented by annually published ship lists (such as the *Annual List of Merchant Vessels of the United States*, a list that was published annually from 1869), which also specified the type and capacity of vessel. More detailed or anecdotal information can sometimes be gleaned from contemporary newspaper accounts, construction plans, and wreck reports filed by agencies, such as the United States Life Saving Service. In other words, basic descriptions are available for most vessels, along with highly specific details for some.

A mainstay of normal efforts to identify wrecks is to rely on such highly detailed historical information whenever possible. The caveat in the use of historical records is that an unknown proportion of the wooden vessels operating on the lakes were *not* registered, and as such little can be said regarding their design, size, or age. Furthermore, in the case of broken or scattered wrecks, many of the highly detailed aspects of vessel design and configuration will be unidentifiable in the archaeological context. It is also true that historical sources do not always agree. In the present applications, the normal processes of source criticism and evaluation are applied, but in more ambiguous or contentious cases, a probabilistic treatment of the historical sources might be pursued.

On the archaeological side of the ledger, there are both pluses and minuses. By their very nature, broken or scattered wrecks are not intact vessels, but rather are bits and pieces of a vessel, in varying states of coherence. In some cases, large subunits of the vessel may remain intact, as in a segment of hull, or the bilge structure. In other instances, only loose timbers and fasteners will be observed. Fortunately, given the canons of vessel construction, it is often possible to estimate the overall size and capacity of the ship based on the dimensions of specific timbers and fasteners, although this process is not without its potential pitfalls (see discussion of vessel capacity below). The major limiting factor for estimating the characteristics of a broken wreck is which pieces remain in situ to be identified, and the degree to which subunits of the vessel remain intact and can be associated with one another.

This latter point alludes to the other difficulty in working with scattered wrecks. As wooden vessels come apart after wrecking, pieces of the boat, including large intact subunits, drift away from the wreck site and come to be deposited in debris traps, which may be at a significant distance from the wreck site. These locations often contain debris from more than a single wreck (for a fuller discussion of this process, see Refs. [23,29]). For any archaeological assessment, therefore, a determination of whether the find represents a wreck site or a debris trap is the necessary first step.

When the range of historical and archaeological factors was considered, five general features were selected for the initial model. These features include: two relating to vessel size (maximum observed length and estimated gross tonnage), the propulsion system (distinguishing steam from wind driven vessels), the cargo, and the specific location of the identified wreck site. When examining an intact vessel, most if not all of these features are unambiguous. In the context of a scattered wreck site, the determinations themselves are probabilistic statements, depending on the particular pieces of evidence preserved at the site. As such, the model necessarily incorporates a second level of Bayesian evaluation, in order to reflect the certainty or confidence that could be placed on the archaeological assessment of each of the five categories of evidence to be evaluated (see below).

As a starting point for evaluating historical identifications, the very general categories of evidence employed contrast starkly with the highly detailed information normally employed by wreck historians. This reflects a fundamental difference in the evaluation of broken as opposed to intact wrecks. In the second test example, the *Bayes Boats 1* probabilities are directly compared with the results of precisely this kind of detailed historical research.

The basic structure of the *Bayes Boats 1* evaluation is illustrated schematically in Fig. 1. The main evaluative process runs from left to right across the figure, starting with the initialization of the alternative hypotheses (vessels) with equal prior probabilities and ending with the final posterior probabilities for these hypotheses, stepping through each of the five evaluations in turn. In each of the test 'boxes' the documented characteristics

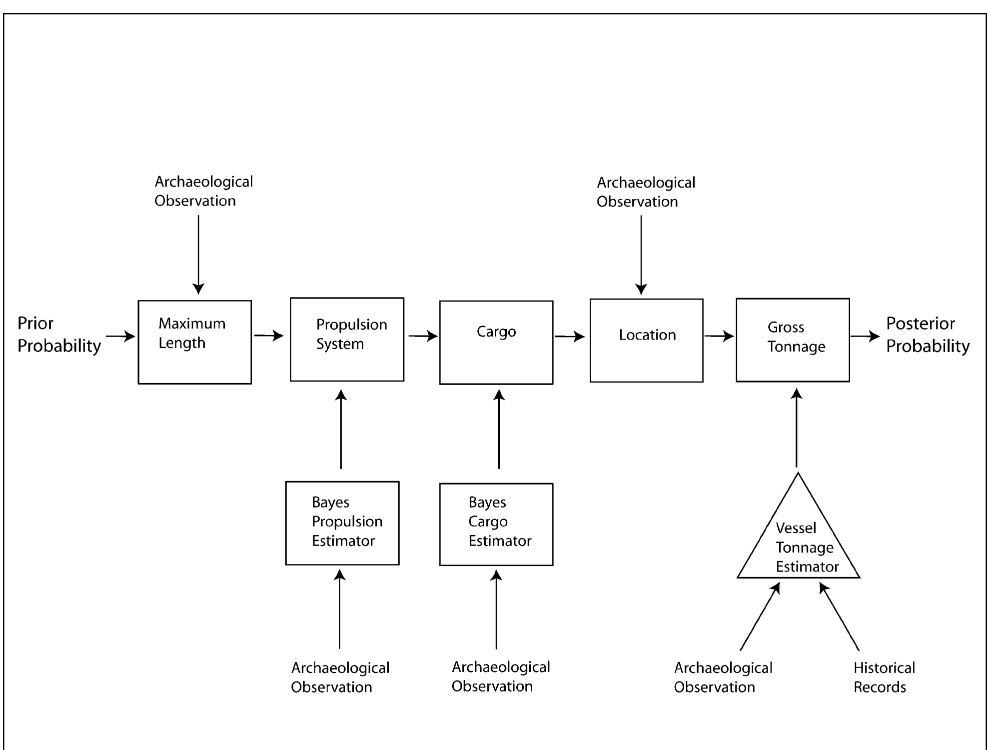

Fig. 1. Schematic representation of the *Bayes Boats 1* system. The model runs from left to right, beginning with prior probabilities and ending with final posterior probabilities for each historic vessel considered. Rectangular boxes represent Bayesian evaluation steps, while the triangle represents a non-Bayesian evaluation. Arrows represent the flow of information into the main evaluative system.

for each historic vessel are compared against the observed archaeological value from the wreck site. The resulting probability reflects the updated probability that, given the archaeological evidence, the historic vessel in question is the source of the wreck. This main evaluative level of the *Bayes Boats 1* system was implemented as a series of macros in Microsoft Excel.

Before going further, several properties of the evaluative model must be clarified. The first concerns the order in which the different classes of archaeological evidence are considered. While the particular order represented in the model makes logical sense, it has no effect on the posterior values generated. The posterior probabilities remain the same regardless of the order in which the tests are evaluated.

The second feature of this model is that the posterior probabilities associated with each vessel are independent of the values for any other vessel. Said another way, the posterior probability derived for one vessel identification is not conditional on the probability associated with any other vessel. While the initial prior probabilities are distributed evenly among the set of plausible historical identification hypotheses, in the evaluation process, each potential identification is contrasted with a 'not this vessel' hypothesis, rather than with a head-to-head comparison between the other plausible vessels. As such, the set of posterior probabilities provides an assessment of how strong each identification is given the available data and a relative sense of the strength of the other plausible identifications. This approach is explicitly open ended, allowing for the potential that the identity of the wreck might be 'none of the above'. Since we cannot be certain that the specified set of plausible wrecks is exhaustive, the recognition of such outcomes is particularly important.

A final issue concerns the conditional probability values assigned in each of the five test areas. Given the mechanics of calculating Bayesian probabilities, extreme values (those approaching either 0.0 or 1.0) exert a disproportionate effect on the posterior probabilities. It might be argued that this is as it should be, since these values tend to irrevocably support or undercut particular hypotheses. Yet, given the uncertainties of archaeological observation in scattered wreck sites, as well as ambiguities in the actual details of vessel construction and historical records, it did not seem desirable to let any single category of information overwhelm the entire evaluation process. Since in Bayesian calculations, a probability value of $P = 0.5$ neither supports nor rejects a given hypothesis, the maximum conditional probability for assessment was set to a value of $P = 0.75$, which represents strong, but not overwhelming, support for an identification. Likewise, the conditional probability for a given outcome was never allowed to drop to zero. In effect, this assures that each of the five categories of comparison will have an influence on the posterior probability.

While the calculation of posterior from prior probabilities is unproblematic using Bayes' Theorem, the derivation of the conditional probabilities associated with the observed archaeological evidence is more complicated. For example, if a potential vessel was a steamer and the wreck was probably a steamer, the likelihood that the identification is correct increases, while if the wreck were sail driven, then the likelihood would decrease. Unfortunately, it often is not possible to say with certainty whether the wreck was sail or steam driven. Instead, it is necessary to specify *how likely*, given the array of archaeological evidence available, it is that the wreck was sail or steam driven.

For each of the five test criteria, it was necessary to systematically assess the available archaeological evidence. These assessments produced the conditional probabilities required for the main Bayesian evaluations. The nature of the rules that governed the generation of these conditional probabilities varied by test and should themselves be considered first approximations. It will also be noted that some of these probabilities derive from relatively simple examination of the archaeological evidence, while others incorporate their own Bayesian evaluation based on the presence or absence of differing classes of evidence at the archaeological site (Bayesian evaluations in Fig. 1 are represented by rectangles, while the triangle represents a non-Bayesian evaluation, see below). The rules governing the determination of conditional probabilities for each test are briefly summarized below.

3.1. Maximum vessel length

Maximum vessel length compares the maximum observed length of the wreck to the stated length of the historic vessel as given in the ship's registry documents. If the observed archaeological length is less than or equal to the registry length, a probability of 0.75 is assigned. If the archaeological length is greater than the registry length, a probability of 0.25 is assigned to the identification.

3.2. Location

Location compares the historical evidence for the precise place or locality of a historic wreck with the observed location of the wreck site. If the historic wreck can be tied to the immediate vicinity of the archaeological site, a probability of 0.75 is assigned and if not, a probability of 0.25 is assigned to the identification.

3.3. Gross tonnage

Gross tonnage reflects the overall size of the vessel and represents a value that is expressed, to a greater or lesser degree, in all the structural elements of the ship.

Table 1
Estimates of gross tonnage from scantling tables for seven archaeological wrecks from the Pictured Rocks National Lakeshore [15]

Vessel	Year of construction	Official capacity	Scantlings		
			Lake Underwriters 1866	Lloyd's	Bureau of American Shipping
Mary Scott	1866	243	267.0	311.0	354.0
Michael Groh	1867	174	175.0	170.0	250.0
George	1873	790	355.0	469.0	325.0
Mary Jarecki	1871	646	366.0	388.0	418.0
Sitka	1887	1741	900.0	1221.0	992.0
Gale Staples	1888	1847	1094.0	1200.0	1350.0
Union	1861	341	169.0	567.0	250.0

Scantling tables, published by various lake shipping insurers (cf. Ref. [3]), prescribe the width, thickness, and length of most timbers on a vessel based on its tonnage. The size of fasteners, chains, anchors, and myriad other details were similarly prescribed. Such information is of particular value in the case of scattered wrecks since it allows estimates of vessel capacity to be derived from fragmentary remains.

There are, however, several difficulties in the use of insurance scantling tables to estimate vessel tonnage. The published scantlings represent an ideal standard that had to be achieved in order to receive the highest insurance rating. Yet, many lake vessels never met this standard. It was also common for a vessel to be underbuilt (in terms of insurance standards) on one structural element and to be overbuilt elsewhere to compensate. Contemporary publications openly discuss this activity and treat it as an acceptable practice [9]. Evidence of this practice is manifest in the archaeological cases by relatively high coefficients of variation for the estimates of vessel capacity.

It should also be remembered that ship building on the Great Lakes during the nineteenth century was more an evolving craft than an established science, and as such, experimentation and novelty in design were common. This is particularly true in the design of larger wooden vessels, which were fighting an ultimately doomed battle to compete with the large iron and steel boats that increasingly dominated lake commerce [20,26]. Since the size of vessels, and the techniques used to construct them, were evolving, the scantling values also changed over the course of time. As such, a vessel built to standards in 1860 might not match scantling specifications for 1890. In future models, it might be possible to use these changing standards as test evidence. Since vessels were often rebuilt or retrofitted to meet new underwriter standards, such retrofitting might itself provide evidence for the age of the wreck which would in turn support or undermine potential historical identifications. To further complicate matters, the contemporary formulae for calculating registry tonnage varied between nations and changed over time [19].

Ideally, scantling tables contemporary with the construction of each vessel would be considered but, since the precise age of construction is often not obvious from a scattered wreck, some other approach to handling this source of variation had to be devised. To better understand how to handle this ambiguity, a series of preliminary tests were conducted to assess the predictive value of various timbers from lake vessels and differing published scantling guides. These results are summarized in Tables 1 and 2.

Seven vessels reported from the Pictured Rocks National Lakeshore [15] provided a set of cases where estimates of capacity based on archaeological remains could be contrasted with the registered capacities of the vessels. Table 1 presents the registered capacity of each boat, along with the estimate of capacity that is derived using three standard scantling tables – the 1866 Board of Lake Underwriters table [3], the Lloyd's table of 1900, and the current table of the Bureau of American Shipping [9].

Table 2
Combined estimates of gross tonnage for seven archaeological wrecks from the Pictured Rocks National Lakeshore

Vessel	Year of construction	Official capacity	Combined estimate	Standard deviation	Coefficient of variation (%)	Within 1 STD	Within 1.5 STD
Mary Scott	1866	243	389.3	163.8	42	Yes	Yes
Michael Groh	1867	174	312.5	130.1	42	No	Yes
George	1873	790	547.9	347.1	63	Yes	Yes
Mary Jarecki	1871	502/646*	444.3	256.1	58	Yes	Yes
Sitka	1887	1741	1190	415.5	35	No	Yes
Gale Staples	1888	1847	1306.3	626.9	48	Yes	Yes
Union	1861	341	462.5	165.6	36	Yes	Yes

* Reflects change in rated capacity after rebuilding.

Several important trends are visible from the table. The first is that the estimates yielded by the three scantling tables are similar, but not identical. In general, the estimated tonnages generated from the 1866 scantling table are lower than both the Lloyd's or Bureau of American Shipping estimates. Secondly, all of the scantling tables tend to underestimate the capacity of the five largest lake vessels in the sample.

In light of these findings, it can be expected that estimates of vessel capacity based on the remains of scattered wrecks will be affected by the date of the vessel's construction, by the absolute size of the ship in the case of very large vessels, and by a general tendency of the scantling tables to underestimate vessel capacity, and of larger vessels in particular. It is also important to note that not all timbers are equally diagnostic of tonnage. As such, the particular timbers that can be observed at the site will affect the precision of the archaeological estimate.

In an effort to compensate for this predictable underestimation of vessel capacity, all three of the scantling tables are employed (since there is no a priori basis for knowing the precise age of the wreck) with the tonnage value assigned to each observed timber representing the largest capacity associated with that timber across the three tables. If the dimensions of a timber or fastener are associated with a range of capacities, the midpoint of the range is selected. In cases where the range is open-ended (e.g. >200 tons) the largest specific value derived from the other timbers plus 200 tons is used as the upper end point for determining the midpoint of the range. All of the observed tonnage estimate values derived from the wreck site are then averaged in order to produce a mean estimate for the wreck and an associated standard deviation.

Mean capacity values and standard deviations were calculated for the seven Pictured Rocks wrecks and evaluated against the registered capacities of the associated historical vessels in the sample (Table 2). When a span of plus or minus one standard deviation is considered, the estimates include the registered value in five of the seven cases (71%), and all of the cases are correctly included if a span of plus or minus 1.5 standard deviations is utilized. The coefficients of variation that result from these estimates are high, and suggest that considerable deviation from the norm is present among the differing timbers used in the construction of the individual boats.

Using these figures as a guide, the vessel tonnage estimator assigns probabilities to each vessel as listed in Table 3.

Table 3
Assigned conditional probabilities for estimated wreck gross tonnage

Vessel capacity	Assigned probability
within 1 standard deviation	0.75
within 1.5 standard deviations	0.50
within 2 standard deviations	0.25
outside 2 standard deviations	0.05

3.4. Propulsion system

While determining the primary means of propulsion is usually obvious for an intact vessel, the answer becomes more ambiguous in the case of broken up wrecks. Some vessel features can be decisive, but when they are not observed, or when they are observed together with structures associated with the opposing propulsion method, the answer becomes less clear. The archaeological problem is exacerbated by the fact that some Great Lakes vessels employed both steam and sail, others were built for sail and then converted to steam, and yet others were built for steam and then converted back for propulsion by wind and hawser line.

To cope with these uncertainties, the *Bayes Boats 1* system employs a second level of Bayesian evaluation to derive the conditional probability for wreck propulsion. For this evaluation (and the evaluation of cargo, see below) a Bayesian Belief Network was constructed using *MSBNx*, a program distributed as freeware from Microsoft [14] (for the graphical representation of these networks in the figures below, a commercial Bayesian network program [*Netica* (Norsys Software Corp)] has been used). This program allows for a series of relationships to be constructed graphically and provides the estimated probabilities for differing system states given specified 'known' values for some of the variables. A particularly useful aspect of the program is that variables can be defined as having one of a number of 'known' states or as being simply unobserved. In the present module, for example, the archaeological observation state for a ship's propeller may be 'yes' (present and observed), 'no' (not present) or 'unobserved' (not sure if it was there or not). As such, if a wreck site contained only the forward portion of the vessel, the value for propeller should properly be coded as unobserved rather than not present. Another useful aspect of the Bayesian network approach is that given a set of relationships, it provides a numerical breakdown showing which single piece of new information would have the greatest impact on the posterior probabilities. While this feature is not utilized in the present application, it is easy to imagine how such a system might be used to guide future field strategies.

The schematic diagram of the propulsion probability estimator is presented in Fig. 2. Starting at the top of the diagram, the program assumes there is an equal probability that the original vessel was steam or wind driven (towed barges are treated as wind driven, while vessels employing mixed steam and sail propulsion are treated as steam driven). The system then evaluates the probability that a series of five distinct vessel features will be present if the vessel is steam driven, and likewise if it is sail driven. The intermediate probability values associated with the wreck site of a steamer and a schooner are presented in Figs. 3 and 4, respectively. Some

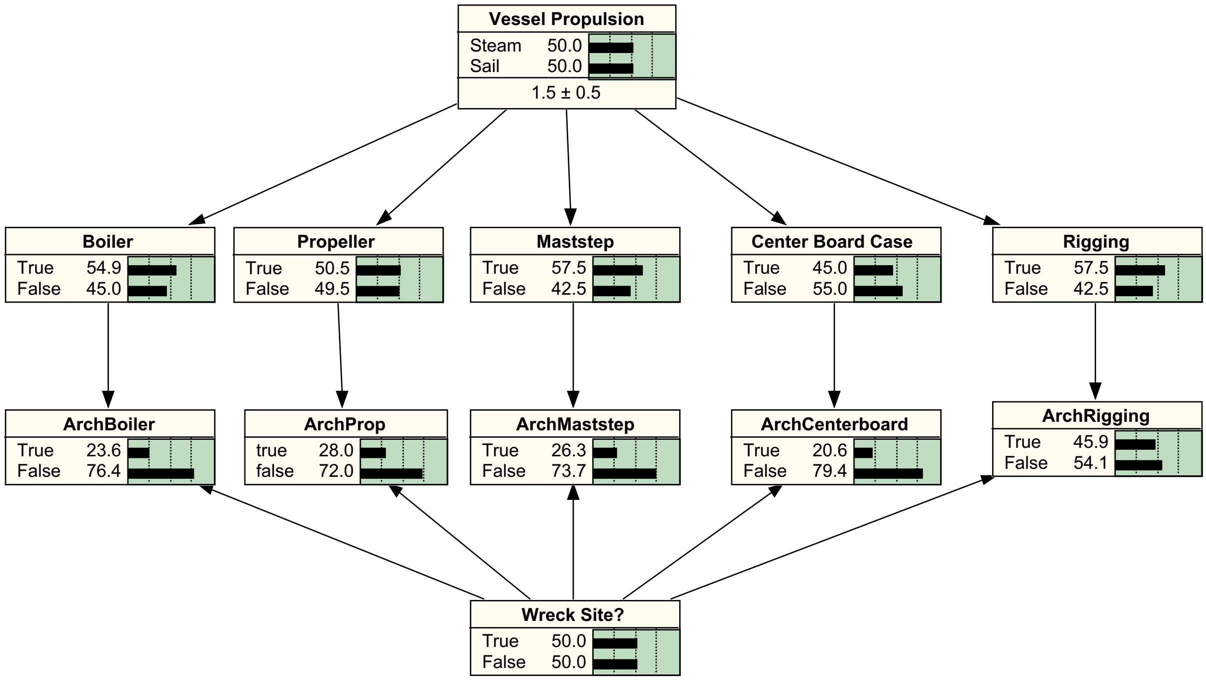

Fig. 2. Bayesian Propulsion Network with no evidence supplied.

features are highly diagnostic, such as the existence of a propeller (or paddle wheel) or the steam boiler apparatus, while elements such as a mast step or sail rigging are more ambiguous. After establishing the likelihood that these features would be found with a sail or steam boat, respectively, the program then estimates the probability that each element would be present, if the site represents a true wreck site or if it is a secondary debris accumulation. With this information, the program estimates the probability that each feature will or will not appear in the archaeological assemblage. It is then simply a matter of filling in the blanks to indicate what kind of site it is, and which features are, or are not, observed. Based on these observations, the program calculates the conditional probabilities that the archaeological wreck was steam propelled or wind driven. The probability of

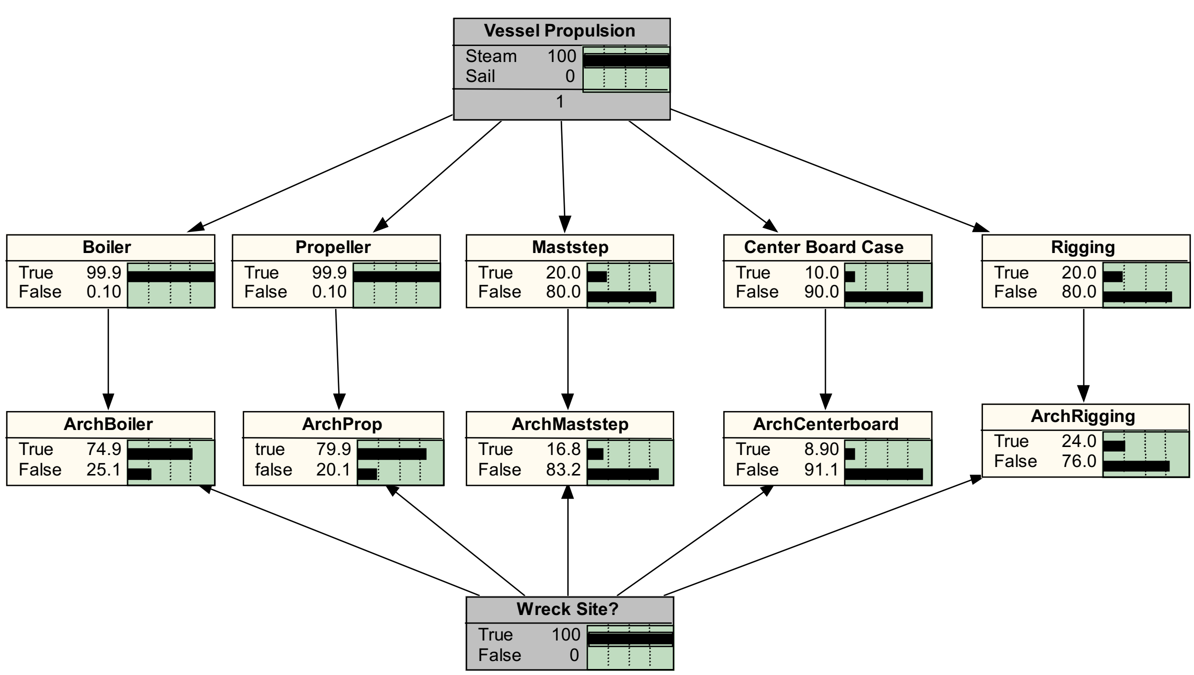

Fig. 3. Bayesian Propulsion Network values for the evidence of a steamer at a wreck site. Determined values are shown in shaded boxes.

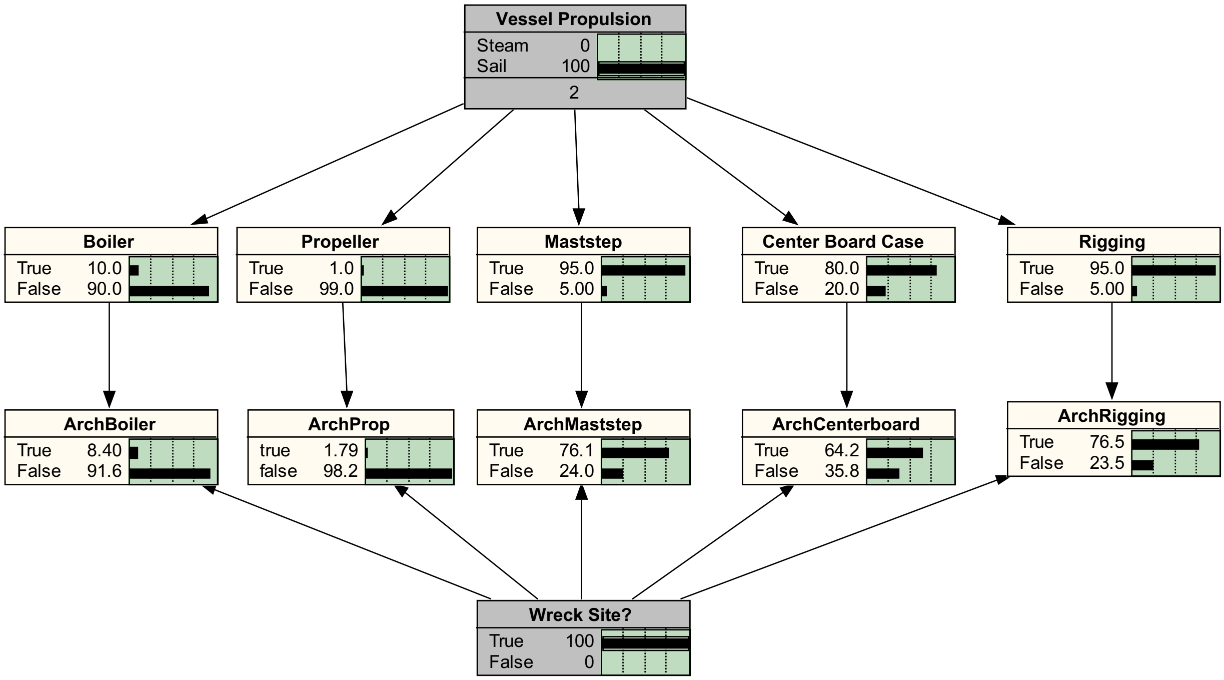

Fig. 4. Bayesian Propulsion Network values for the evidence of a schooner at a wreck site. Determined values are shown in shaded boxes.

sail and the probability of steam are both multiplied by 0.75 (reflecting the weighting of the propulsion evidence in the overall evaluation as previously noted) and the appropriate conditional probability is assigned to each historic boat depending on its reported method of propulsion.

3.5. Cargo

Vessel cargo, like the system of propulsion, is a category of information that is susceptible to a number of intervening depositional and post-depositional factors. Great Lakes vessels carried a variety of cargoes from ores and stone to wheat, timber, and coal. They also occasionally traveled in ballast with no cargo. To compare vessels with known cargoes to the deposits at a scattered wreck site, it is necessary to account for the buoyancy of the cargo (will it sink like stone, or simply float away like timber) and for the relative likelihood that the cargo would be salvaged after wrecking?

The Bayesian cargo estimator, like the propulsion module already described, utilized *MSBNx* to associate the likelihood of differing cargoes being present given a set of archaeological observations. The schematic diagram of the relationships is presented in Fig. 5. While a great deal of detail might be incorporated into such a model, for the present purposes several simplifications were introduced. The expected Great Lakes cargo was reduced to six categories: ores, stone, brick, timber, coal and ballast. The category 'in ballast' was interpreted to include not only vessels that were truly in ballast, but also passenger ships and vessels that carried a mix of small goods. In the present samples, other buoyant perishable cargoes such as wheat or cotton were not encountered. The pattern of post-wreck distribution of these materials would be similar to that of lumber. The model also simplified the distribution properties of the cargo into three possible states – abundant/in hull, scattered, and rare – to reflect the character of the cargo's distribution at the site.

The model starts with an equal probability assigned to each of the potential cargo type hypotheses. It then calculates the probability that remains of the cargo will be found at the archaeological site, depending on whether the location is a wreck or accumulation site and on the density of material. It is also necessary to ascertain whether the vessel was steam or wind driven, since a steamer is expected to carry some coal within its hold, regardless of its primary cargo (while some sailing vessels carried smaller quantities of coal to run steam driven deck machinery, the quantity that would be expected to survive at a wreck site is negligible). This value is derived from the propulsion evaluation already described, with the most probable propulsion method being assigned to the wreck.

With these values specified, the model provides a conditional probability associating each cargo type with the observed archaeological remains. Each probability is multiplied by 0.75 (for the reasons described earlier) and is applied in the cargo module of the main analysis. For cargoes with estimated probabilities less than 0.067, the probability 0.05 is used in the main analysis.

Using the probabilities generated by these differing modules, the main program systematically compares the

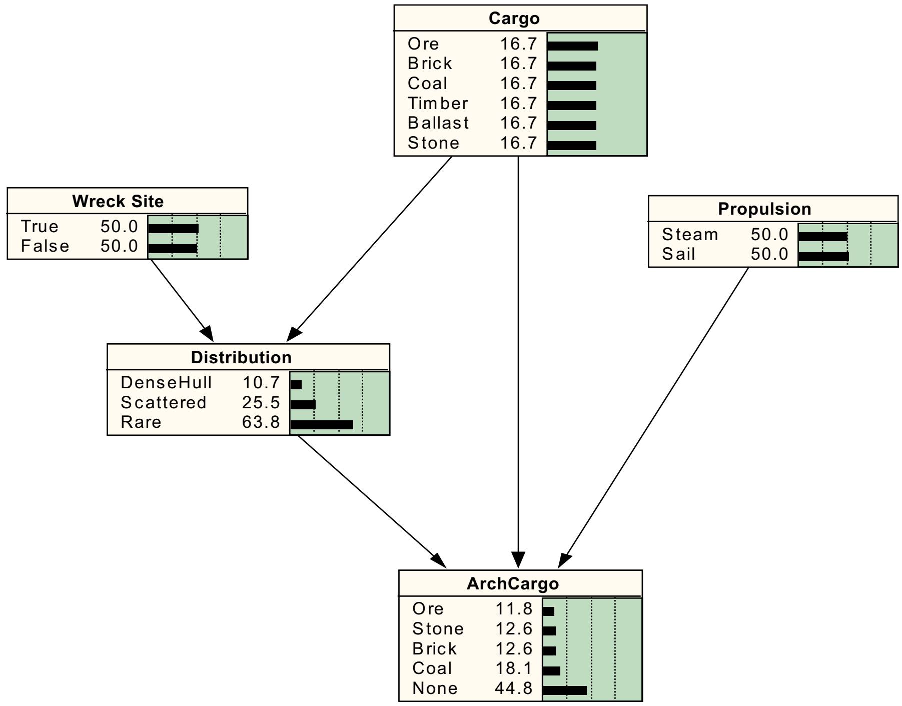

Fig. 5. Bayesian Cargo Network with no evidence supplied.

observations derived from the wreck site with the known characteristics for each historic vessel. The probability of each historic vessel being linked to the wreck is updated at each step and ends with the final posterior probability for that vessel. The working of the system will be illustrated in the two archaeological tests that follow.

4. Archaeological applications

To evaluate the performance of the *Bayes Boats 1* system two test cases were selected from the Great Lakes region (Fig. 6). The first evaluates the identification of a single wreck in the Au Sable Shores region of western Lake Huron from among a set of ten vessels whose loss is attributable to the area. This test is designed to illustrate how the model works and particularly to illustrate how the relative probability associated with each identification is modified as each new piece of evidence is evaluated.

The second case looks at a series of archaeological wrecks that have been reported from the Pictured Rocks National Lakeshore on the southern shore of Lake Superior. In the case of these wrecks, the accepted identifications have been developed via intensive historical research by a leading authority in Great Lakes maritime history. This test is designed to show how the *Bayes Boats 1* system results compare with these standard historical identifications.

4.1. A keelson at Au Sable Point

The long keelson and bilge assembly at Au Sable Point (20UH550) was first recorded in 1998, and has been the subject of archaeological monitoring continuously since that time, although no excavation or modification of the wreck site has been undertaken [24]. The fragment is oriented roughly north south and rests at an angle to the present day beach (Fig. 7).

The fragment represents a keelson with sister keelsons on either side, all overlain by a rider keelson. Sixteen sets of floor frames were observed to protrude from beneath the keelson assembly and it is supposed that these overlie the vessel's keel. The complete keelson assembly is 30 inches wide, with an observed length of at least 142 feet, although both ends of the assembly remain buried (note: all vessel measurements are reported in terms of the 'English' system of measurement to facilitate comparison with contemporary construction plans and scantling tables).

The timbers used in the construction of the keelsons all appear to be a similar 10″ thick, but with the sister keelsons being 9.5″ in width, and the rider (and presumably central) keelson being 11″ wide. The timbers were fastened with parallel rows of 1″ drift bolts, two per member, and spaced at intervals of 22″ along the length of the keelson. An extra pair of bolts is found at the beginning of each scarph joint. Scarphs on the keelson assembly were 60″ long and were secured by a series of three 1″ diameter through bolts, spaced at

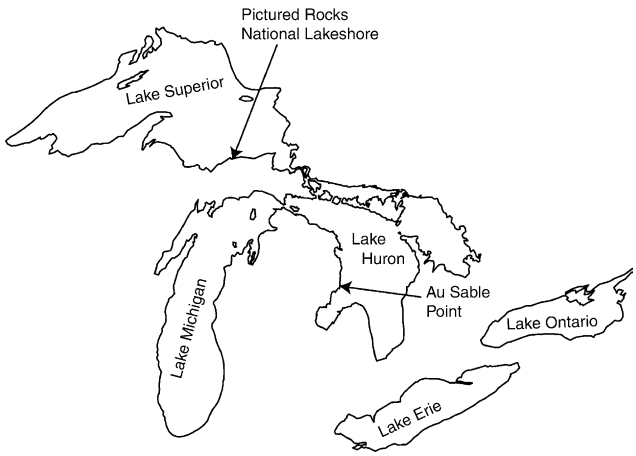

Fig. 6. Map of the Great Lakes region, showing the location of the two test cases described in the text.

intervals of 20″. The joins appear to be simple flat scarphs. It is interesting to note that the break in the keelson assembly is located at a point where scarphs in the rider keelson and a sister keelson coincided.

The frames that protrude from beneath the keelson assembly are 8″ paired, with a thickness of 11″ as they emerge from beneath the keelson (20″ from the centerline) and a thickness of 9″ at a point 80″ out from the centerline. An upturn in the shape of the frames is first visible about 53″ out from the centerline. The frames line up with the visible drift bolts on the upper keelson elements and confirm that the vessel had 22″ centers.

During the period that the keelson assembly has been observed, additional elements have been recorded including a mast step, a possible second mast step, and a structure that appears to represent the centerboard case of the vessel. The mast step was situated on top of the keelson assembly and was a total of 80″ long. The step was held in place by 7/8″ drift bolts and three 7/8″ transverse through bolts. The opening for the mast base was 17″ by 13″.

A final point regarding the archaeological site is that while the Au Sable Point locality has been identified as a debris trap and contains the remains of several vessels, the particular archaeological remains in question do, in all probability, represent an actual wreck site. This is because the embedded bilge structure of the boat is the least likely portion of the wreck to drift from the wreck site [23]. Given this, the identification experiment will consider all of the timbers directly attached to the keelson assembly but will not incorporate other free pieces of wreckage in the general vicinity of the keelson.

A number of historic wrecks are attributable to the Au Sable Point locality that might conceivably be linked with this site. These vessels and their basic attributes are summarized in Table 4. Among these is the winter wrecking of the schooner barges *Mears* and *Midnight* whose breakup on Au Sable Point is dramatically recounted in the Annual Report of the United States Life Saving Service of 1890 [1]. These two vessels are generally accepted as the most likely attribution for the Au Sable Point wreckage [22,24]. They also present their own problems for the analysis since they represent two very similar ships that were wrecked at the same time within a matter of 100 m or less of one another.

The baseline data for the wreck are that it has a length of at least 142 feet and an estimated capacity of 349 gross ton. Table 5 presents the conditional probability values generated for the wreckage. The second column presents the length, capacity estimate, and standard deviation of the capacity estimates, along with the probability estimates generated by the Bayesian network models for the two possible propulsion systems and the differing possible cargoes. The third column

Fig. 7. Partially exposed keelson assembly at Au Sable Point, Michigan.

Table 4
Characteristics of historic vessels lost in the Au Sable Point region

Vessel	Length	Capacity	Cargo	Propulsion	Location
Mears	172	429	Lumber	Sail	Probable
Midnight	136	382	Lumber	Sail	Probable
Banner	140	289	Lumber	Sail	Improbable
Hercules	183	559	None	Steam	Probable
Langell Boys	151	387	None	Steam	Improbable
John Shaw	206	928	Coal	Sail	Improbable
Thomas Sheldon	194	669		Sail	Improbable
Summit	126	226	Ore	Sail	Improbable
Volunteer	50	31	Lumber	Sail	Probable
Wayne	56	80		Sail	Probable

presents the conditional probabilities that are associated with each of the five observation sets in the main *Bayes Boats 1* system. These conditional probability values as applied to each of the plausible identifications are presented in Table 6. It is worth noting that the second column of Table 5 provides what amounts to a probabilistic description of the vessel that produced the wreck. In the absence of a good fit between the wreck site and a known vessel, this description could itself become an important tool in guiding further historical research.

Since only a single wreck is considered in this example, it is easy to see how the overall evaluation process operates. Take, for example, the probabilities associated with the propulsion system (Fig. 8). Given

Table 5
Archaeological estimates and conditional probability values generated from the Au Sable Point wreck site

Evidence category	Archaeological observation or estimated probability	Conditional probability (P value)
Vessel dimensions		
Maximum length	142 feet	Reg. length GE 142 (0.75) Reg. length LT 142 (0.25)
Mean vessel capacity	349 gross tons	253.8–444.2 (0.75) 206.2–491.8 (0.50) 158.6–539.4 (0.25) LT 158.6 or GT 539.4 (0.05)
Standard deviation of capacity estimate	95.2 tons	
Propulsion		
Probability steam driven	0.050	0.037
Probability sail driven	0.950	0.713
Cargo		
Probability ore	0.085	0.064
Probability stone	0.045	0.034
Probability brick	0.045	0.034
Probability coal	0.092	0.069
Probability lumber	0.366	0.275
Probability ballast	0.366	0.275

the available archaeological evidence, which included the presence of mast steps and a centerboard case and no observation of piping or other elements of the boiler system, the Bayesian Propulsion Network estimated the probability of sail power at the level of about 95%, as opposed to a 5% probability that the vessel was steam powered (these values can also be found in Table 5). These estimates are the basis for the conditional probabilities used in the propulsion system test (0.7127 for sail, and 0.0373 for steam). These conditional probabilities are assigned to each historical vessel in the sample. As an inspection of Table 6 will show, the two steam powered vessels in the historic sample, the *Langell Boys* and the *Hercules*, both have conditional probabilities under the propulsion category of 0.0373, while the remainder of the vessels are assigned a probability of 0.7127.

The Bayesian network estimates for vessel cargo highlight the effect of weak or ambiguous archaeological evidence. Since no cargo remains were found with the Au Sable Point keelson, considerable ambiguity is present in the probabilities associated with the differing cargoes (Fig. 9), and particularly between the categories of lumber (a commonly transported cargo that would not remain at the wreck site) and a vessel carrying no cargo. While such a result is far from conclusive, it is a realistic assessment of the archaeological evidence. And even as the probabilities are low, the difference in relative value is still of use in distinguishing among the potential historical cases.

While ambiguous archaeological evidence can weaken the overall discriminatory ability of the analysis, missing historical information can artificially inflate or deflate the probability of the vessel with the missing data, relative to all the other vessels being considered. The vessels *Thomas Sheldon* and *Wayne* are a case in point, since their cargoes at the time of wrecking could not be determined from historical records. While a conditional probability value of $P = 0.5$ would be neutral in the strict comparison of the hypotheses '*this* boat' versus '*not* this boat', the resulting support for the vessel would be substantially higher than for any of the vessels with a known cargo. As such, the apparent likelihood of the identification would be artificially inflated due to the absence of evidence. To prevent this undesirable effect, vessels with missing historic documentation are assigned a conditional probability that represents the midpoint of the range of conditional probabilities assigned for the evidence category. This is in keeping with placing the case logically midway between the hypotheses of supporting or refuting the specific vessel [11], while retaining some relative comparability among the full set of plausible identifications.

Once the conditional probabilities have been determined, it is a simple matter to pass the set of plausible historical identifications through the main *Bayes Boats 1*

Table 6
Conditional probabilities for historic vessels given archaeological observations at the Au Sable Point wreck site

Vessel	Length	Capacity	Cargo	Propulsion	Location	Posterior
Mears	0.75	0.75	0.275	0.713	0.75	0.7380
Midnight	0.25	0.75	0.275	0.713	0.75	0.2383
Banner	0.25	0.75	0.275	0.713	0.25	0.0336
Hercules	0.75	0.25	0.275	0.037	0.75	0.0049
Langell Boys	0.75	0.75	0.275	0.037	0.25	0.0049
John Shaw	0.75	0.05	0.069	0.713	0.25	0.0011
Thomas Sheldon	0.75	0.25	0.169	0.713	0.25	0.0184
Summit	0.25	0.5	0.064	0.713	0.25	0.0021
Volunteer	0.25	0.05	0.275	0.713	0.75	0.0055
Wayne	0.25	0.05	0.169	0.713	0.75	0.0030

system. The stepwise sequence from prior probabilities to final posterior probabilities is presented in Figs. 10–15. The posterior probabilities are also provided in the final column of Table 6. As these figures illustrate, potential historical identifications are progressively weeded out until at the end, there is only one strong contender, the schooner barge *Mears*. The only other vessel with any support is the *Midnight*.

From Table 6, it is easy to see how various boats came to be eliminated from serious consideration. In the case of the *Midnight*, for example, the crucial difference was vessel length. The registry length of the *Midnight* was less than the observed length of the keelson assembly, and this fact was decisive, despite the similar cargoes and wrecking circumstances. Other vessels dropped out as a result of their propulsion system (*Langell Boys* and *Hercules*), their cargo (*John Shaw* and *Summit*), or a poor fit with the wreck's estimated capacity. On the issue of vessel capacity, it is also interesting to note that, assuming the wreckage is attributable to the *Mears*, the calculation of gross tonnage from the archaeological remains again underestimated the registry value (estimated 349 tons, registered 429 tons) although the value did fall within the one standard deviation range.

4.2. Wrecks in the Pictured Rocks National Lake Shore, Lake Superior

The Pictured Rocks National Lake Shore provides a more challenging test for the *Bayes Boats 1* system since it incorporates numerous wreck sites stretching along a 50 km stretch of Lake Superior (Fig. 6). For this test, the published report of the wrecks and their attribution to historic vessels is taken as the starting point [15]. All of the archaeological information presented in both text and drawings has been used to characterize the

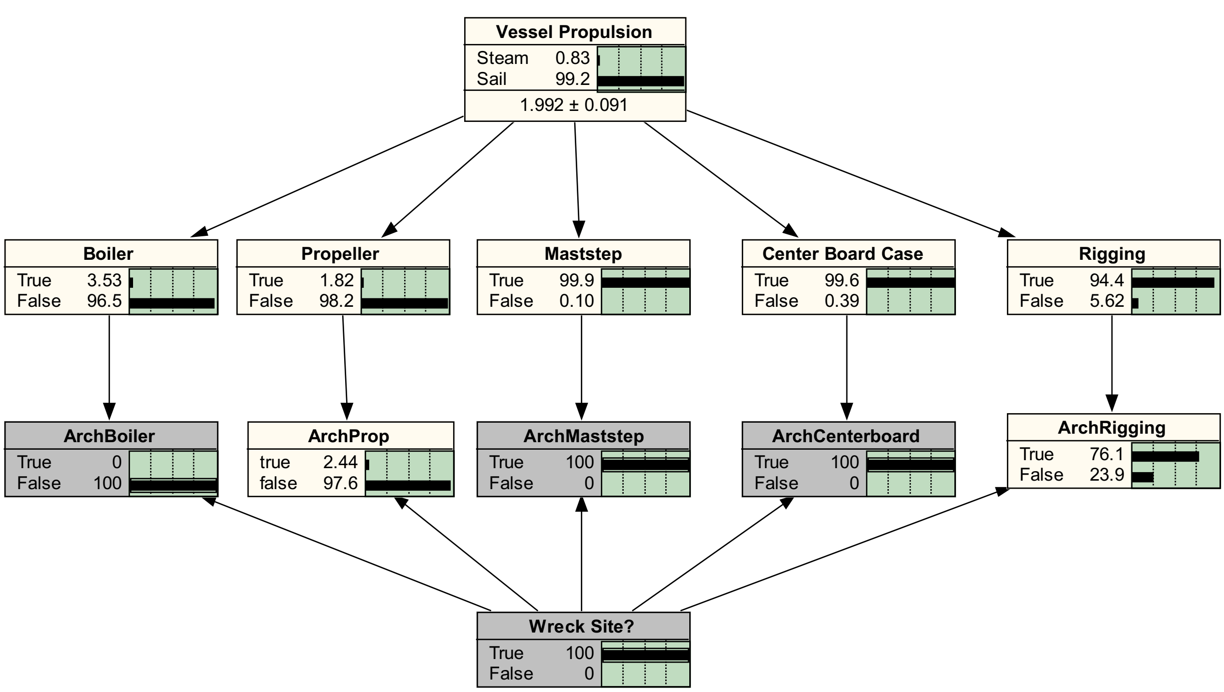

Fig. 8. Bayesian Propulsion Network with data from the Au Sable Point wreck site. Observed archaeological values are shown in shaded boxes.

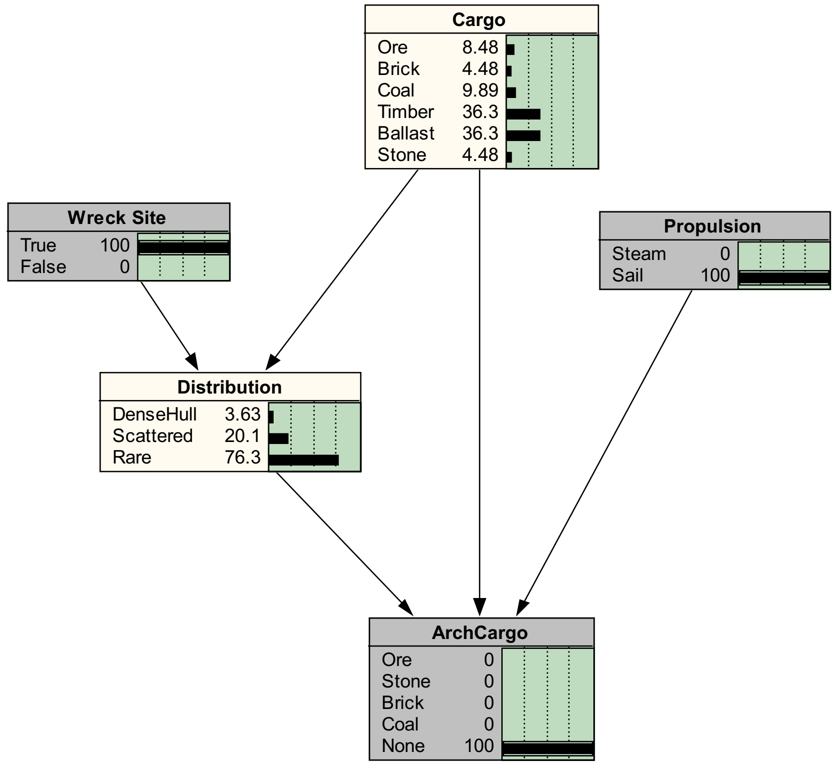

Fig. 9. Bayesian Cargo Network with data from the Au Sable Point wreck site. Observed archaeological values are shown in shaded boxes.

wreck sites, and the universe of likely wrecks developed by Labadie is treated as representing the set of plausible historic wrecks. No new historical or archaeological research beyond that presented in the 1989 report was undertaken for this test.

Several preliminary decisions were necessary in order to apply the *Bayes Boats 1* system to the National Lake Shore sample. First, it was necessary to exclude the intact wrecks from consideration. The intact wrecks are described in an entirely different manner from scattered wrecks in the report, and present a very different kind of identification problem. Attempting to treat them as though they were scattered wrecks would only muddy the comparative value of this test. Vessels that were

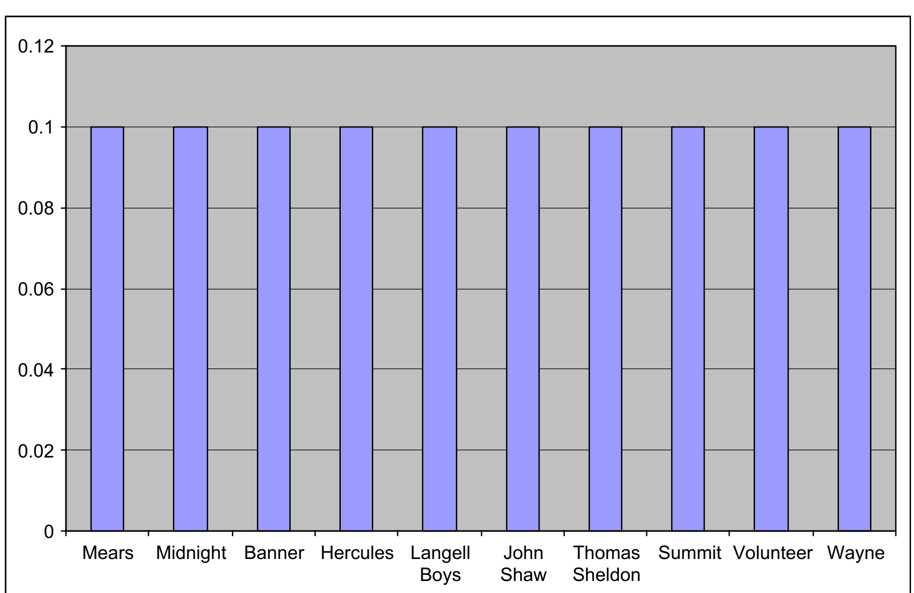

Fig. 10. Prior probabilities for potential Au Sable Point identifications.

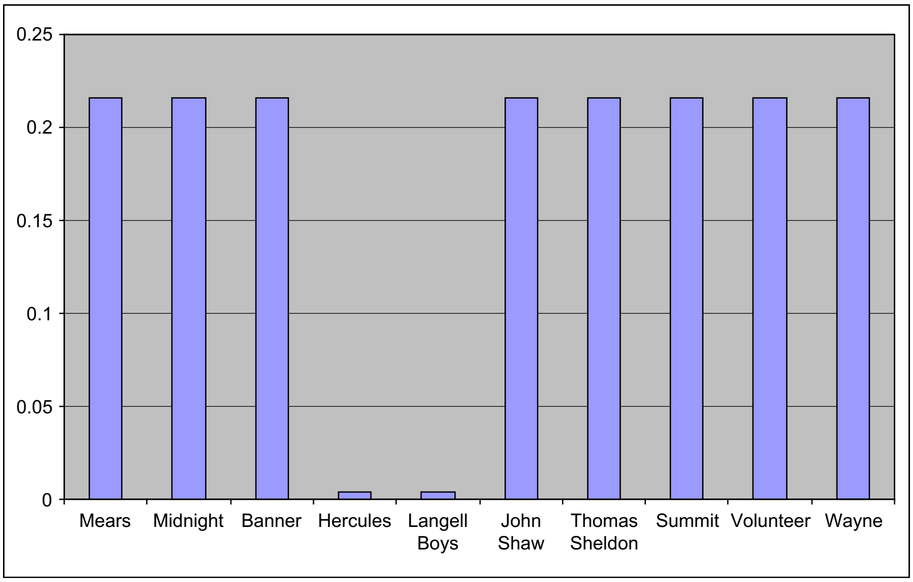

Fig. 11. Probability of identifications after assessment of propulsion.

constructed of steel or iron were similarly excluded from the study.

A second decision made for this study concerns the anecdotal location information assessment. In his description of the wrecks, Labadie identifies a sequence of localities running from west to east along the lakeshore. For the purposes of assessing historic wreck location, historic wrecks attributed to a given locality are treated as supported ($P = 0.75$) for that locality, and not supported ($P = 0.25$) for the other localities.

With these starting conditions, 14 wrecks attributable to six localities were coded. For each wrecked vessel, values for length, estimated capacity, and location were determined, and estimated probabilities for cargo and propulsion system were calculated. These values were then converted into conditional probabilities, as described in the previous example, for comparison with the characteristics for each of the 14 historic vessels. These evaluations resulted in a matrix of posterior probability values for each historic vessel for each wreck site. Descriptions of the 14 historic vessels are presented in Table 7, while the matrix of posterior probabilities is presented in Table 8.

Rather than tracking each vessel individually, it is most useful to consider the overall pattern of outcomes in Table 8. This table can be viewed in two ways. When

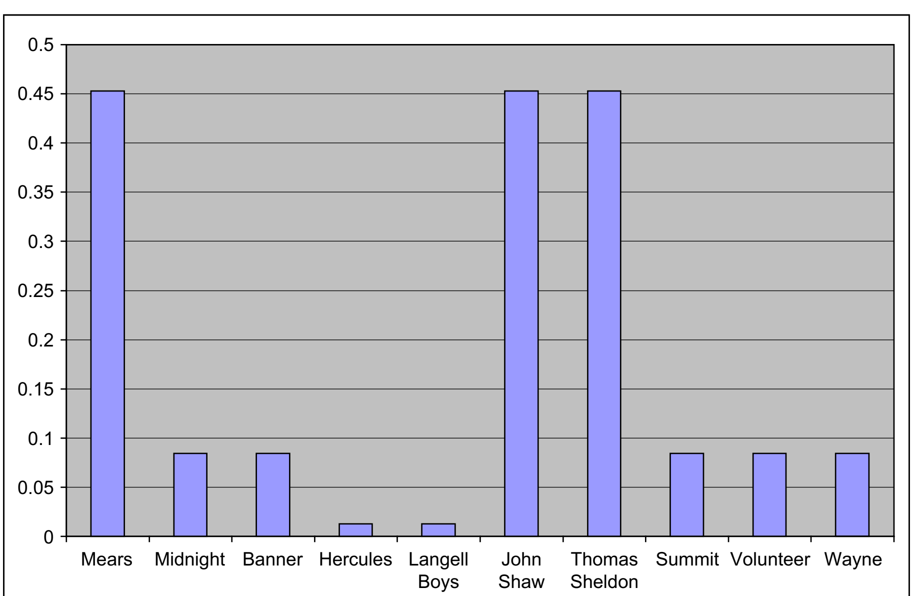

Fig. 12. Probability of identification after assessment of maximum length.

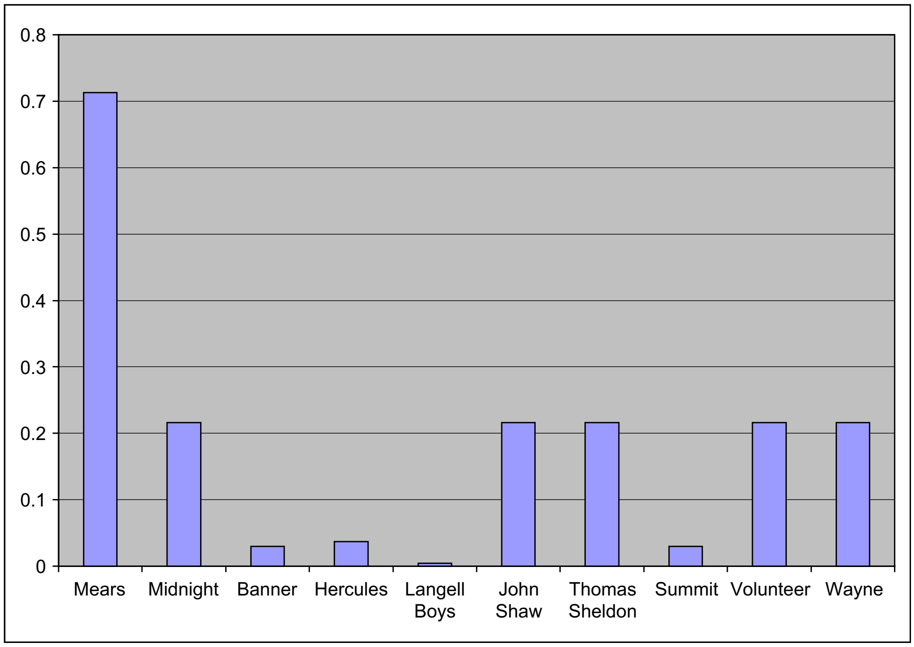

Fig. 13. Probability of identification after assessment of historical evidence for wreck location.

looking down the columns, one is looking for the historical vessel that is the most likely match to the wreckage site. The label on each column represents Labadie's determination of the most likely identification. Looking across the individual rows of the table provides the most likely wreck site for a given historic vessel. These two assessments coincide on the principal diagonal of the matrix; these values are shown in bold in Table 8.

Taking the perspective of wreck sites (columns), it can be seen that there is very good agreement between the most probable historical vessel determined by *Bayes Boats 1* and Labadie's historic identifications. The only mismatch (marked by an asterisk) occurs in the wreckage that Labadie attributes to the *Mary Jarecki*, but which the *Bayes Boats 1* system assigns with equal likelihood to the *Mary Jarecki* and the *Union*. Both vessels receive a high posterior probability ($P = 0.9201$), suggesting that both exhibit a close fit with the archaeological observations. Examination of Table 7 shows that the two boats are similar in cargo, propulsion, length, and

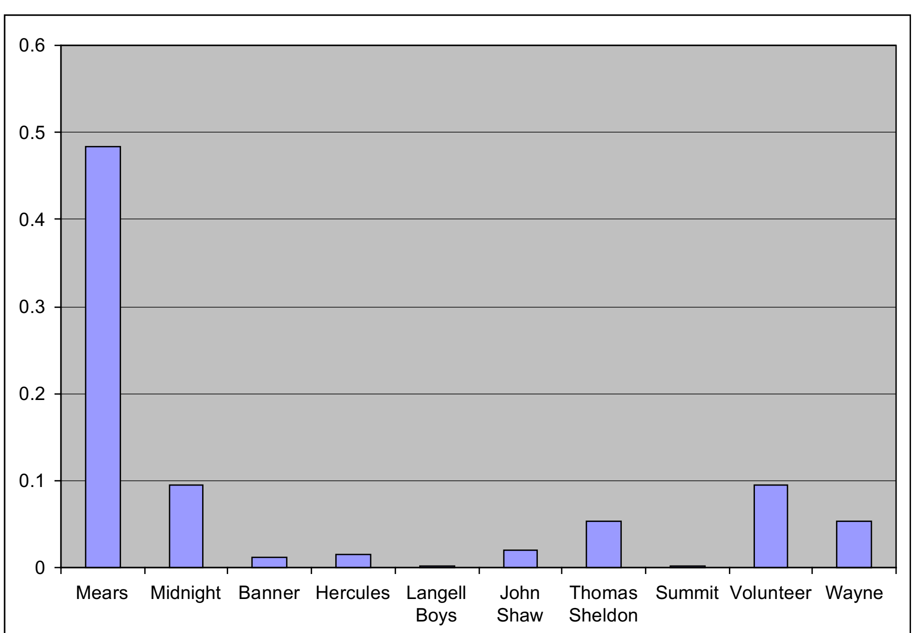

Fig. 14. Probability of identification after assessment of cargo.

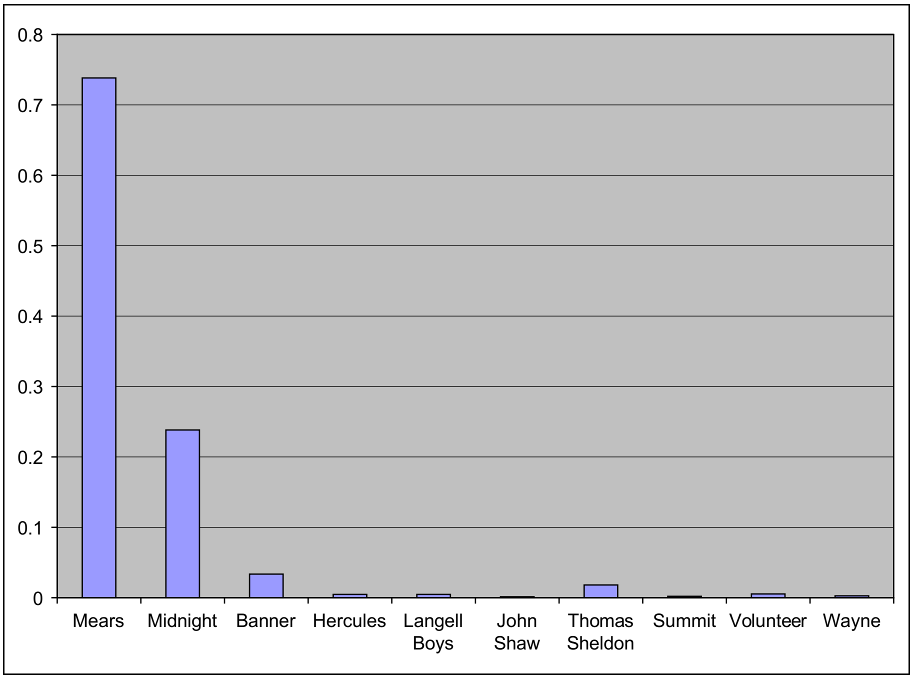

Fig. 15. Final posterior probability of identification after assessment of vessel capacity.

location of wreckage, but the *Jarecki* is almost twice the gross tonnage of the *Union*. In this instance, the capacity estimates derived from the wreck site were exceedingly variable (mean value 444 tons, standard deviation 256). As a result, despite its much larger capacity, the *Jarecki* could not be distinguished from the *Union* given the observable archaeological data.

With the exception of this single tied value, however, the identification with the highest probability associated with each wreck site is the same historical vessel identified by Labadie. Beyond this agreement, though, another important attribute of this table is the reporting of the relative strengths of the potential identifications. For some wreck sites, the relative likelihood value is quite high, such as the values just mentioned for the *Mary Jarecki* and *Union*. In the case of other wrecks, however, even the highest posterior value is quite low. In the case of the supposed *Sitka* wreck site, for example, the highest posterior probability is only 0.1305. This low value is due partially to the low estimate of vessel capacity yielded from the archaeological remains, but also from uncertainty in associating pieces of wreckage from multiple boats that had become commingled in the Hurricane River area. In this sense, the *Bayes Boats 1* model highlights the fact that no historical identification is particularly strong given the current evidence in hand.

If we turn to the rows of Table 8, we can see that there are three instances (probabilities in italic) where

Table 7
Characteristics of historic vessel from Pictured Rocks National Lakeshore

Vessel	Length	Capacity	Cargo	Propulsion	Locality
Bermuda	143	400	Ore	Sail	Munising
Smith Moore	223	1192	Ore	Steam	Munising
Mary Scott	139	243	Ore	Sail	Sand Point
Michael Groh	120	174	Lumber	Steam	Sand Point
Elma	165	401	Lumber	Sail	Sand Point
George	203	790	Coal	Sail	Mosquito Beach
Wabash	140	315	Coal	Sail	Chapel Beach
Superior	191	567	None	Steam	Chapel Beach
Mary Jarecki	180	646	Ore	Steam	Hurricane River
Oneida Chief	127	334	Ore	Sail	Hurricane River
Sitka	272	1741	Ore	Steam	Hurricane River
Gale Staples	277	1847	Coal	Steam	Hurricane River
Union	163	341	Ore	Steam	Hurricane River
South Shore	84	73	None	Steam	Grand Marais

Table 8
Posterior probabilities relating historic vessels to wreck sites in the Pictured Rocks National Lakeshore

Historic boats	Wreck sites													
	Bermuda	Smith Moore	Mary Scott	Michael Groh	Elma	George	Wabash	Superior	Mary Jarecki	Oneida Chief	Sitka	Gale Staples	Union	South Shore
Bermuda	**0.9230**	0.0000	0.5607	0.0008	0.0014	0.0432	0.0082	0.0000	0.0139	0.0801	0.0000	0.0000	0.0006	0.0002
Smith Moore	0.0102	**0.2698**	0.0003	0.0006	0.0001	0.0003	0.0000	0.0139	0.1245	0.0578	0.0476	0.0362	0.0054	0.0162
Mary Scott	0.1290	0.0000	**0.9199**	0.0068	0.0127	0.0432	0.0243	0.0000	0.0016	0.0801	0.0000	0.0000	0.0002	0.0002
Michael Groh	0.0000	0.0004	0.0009	**0.3452**	0.0257	0.0010	0.0005	0.0665	0.0010	0.0018	0.0003	0.0001	0.0146	0.0425
Elma	0.0276	0.0000	0.1967	0.0333	**0.3903**	0.1109	0.0082	0.0002	0.0003	0.0025	0.0000	0.0000	0.0015	0.0005
George	0.0094	0.0000	0.0090	0.0001	0.0016	**0.2789**	0.0144	0.0000	0.0001	0.0025	0.0005	0.0009	0.0002	0.0002
Wabash	0.0031	0.0000	0.0265	0.0007	0.0016	0.0412	**0.8823**	0.0004	0.0003	0.0025	0.0000	0.0000	0.0005	0.0002
Superior	0.0002	0.0076	0.0001	0.0553	0.0087	0.0029	0.0016	**0.3908**	0.0266	0.0018	0.0051	0.0195	0.1176	0.0425
Mary Jarecki	0.0034	0.0015	0.0049	0.0115	0.0016	0.0010	0.0001	0.0139	**0.9201**	0.3558	0.0164	0.1012	0.1289	0.0162
Oneida Chief	0.1290	0.0000	*0.5607*	0.0008	0.0014	0.0432	0.0243	0.0000	0.0139	**0.4395**	0.0000	0.0000	0.0050	0.0002
Sitka	0.0002	0.0135	0.0003	0.0006	0.0001	0.0001	0.0000	0.0139	0.1681	*0.3558*	**0.1305**	0.2525	0.0077	0.0162
Gale Staples	0.0000	0.0120	0.0000	0.0005	0.0001	0.0000	0.0003	0.0123	0.0043	0.0156	0.1235	**0.7083**	0.0073	0.0152
Union	0.0102	0.0001	0.0145	0.0336	0.0047	0.0010	0.0005	0.0139	*0.9201**	0.3558	0.0009	0.0042	**0.3074**	0.0162
South Shore	0.0000	0.0004	0.0000	0.0003	0.0003	0.0000	0.0001	0.0665	0.0001	0.0018	0.0003	0.0001	0.0003	**0.2856**

Column headings represent the most likely vessel attribution for the wreck sites as determined through historical research by Labadie [15]. Values represent posterior probability of the hypothesis 'this boat' versus 'not this boat'. Probabilities in bold reflect a match between archaeological and historical identifications. A probability marked by an asterisk indicates a high value for a wreck with a vessel other than the 'identified' historical boat. Italic probabilities indicate a higher value for a historical vessel than with its 'identified' wreck site.

the most likely wreck site for a given historic boat is not the historically identified wreck site. The difference between this association and the one just discussed is subtle. With the rows, we are starting with a historic vessel and saying which wreck site provides the best match, while with the columns we start from the archaeological site and seek the historic vessel with the greatest likelihood. The *Mary Jarecki* and *Union* have already been discussed. The similarity between the *Oneida Chief* and the *Mary Scott* is clear from Table 7, with the only substantial difference being in the precise location of the wrecks. The third disagreement also involves the *Oneida Chief*; this time it is the historic description of the steamer *Sitka* which has a stronger affinity to the presumptive *Oneida Chief* wreck site. A factor in both of these instances is the very limited archaeological evidence that can be derived from this wreck site. For the presumed *Oneida Chief* wreck site, no estimate of either vessel length or capacity was possible. As such, the assessment had little discriminatory power for any vessel, while the lack of estimate of capacity served to artificially enhance the wreck site's compatibility with *all* of the historic wrecks.

As these results illustrate, attempting to 'find' a particular historic vessel among a set of wreck sites is more affected by archaeological preservation and observation than is the process of determining the source of a wreck site from among a set of plausible vessels. In the latter case, the effects of ambiguous archaeological evidence are spread equally among all potential identifications, while in the former, weak evidence, which neither refutes nor supports a given identification, may still result in relative values that seem comparatively high.

5. Discussion

Overall, the agreement between the identifications produced by the *Bayes Boats 1* evaluation and historical research is quite good. The agreement borders on remarkable when one considers how little detailed historical information the system incorporates, when compared to the extensive scholarly work that resulted in Labadie's historical identifications for the Pictured Rock National Lakeshore wrecks. Aside from doing a good job of linking wreck sites to likely historic vessels, the system also provides a useful general impression of the overall confidence that can be placed in a given historical identification. The association of the *Mears* with the Au Sable Point wreck, for example, is quite strong, while the confidence that can be placed in several of the Pictured Rocks identifications is considerably lower, and may suggest that *any* historical identification is premature in light of existing archaeological information. As such, the system highlights instances where additional targeted archaeological data collection is warranted.

A particularly important, although problematic, aspect of wreck evaluation concerned the estimation of the vessel's gross tonnage. Capacity is a value that can be estimated even from a badly broken up and scattered wreck, but the accuracy and precision of such estimates still leave much to be desired. To the extent that most estimates for larger vessels tend to underestimate the gross tonnage, a revised version of the current procedure might usefully introduce an asymmetrical assessment of the ranges around the estimate. For example, instead of using plus or minus one standard deviation of the estimate, the highest probability band might include plus one standard deviation or minus 0.5 of the standard deviation. A second factor that is not directly reflected in the estimate is the particular set of ship timbers and fasteners that are utilized for making the estimate. Some elements of vessel construction are simply more decisive estimators of capacity than are others. For example, an intact bilge assembly with keelsons, shear strakes, frames and heavy fasteners clearly provide a better estimate of vessel capacity than will a few strakes from the upper hull. Future efforts will be directed at designing a Bayesian Belief Network to assign conditional probabilities over the range of possible capacity values, which will incorporate these kinds of details from the wreck site assemblage.

The issue of missing data arose in both archaeological cases. It seems clear from these tests that missing (or poor quality) data have quite differing effects, depending on whether they occur on the historical or the archaeological side of the evaluation. In the case of poor archaeological evidence, as in the Pictured Rocks example, the main result is an inability to distinguish between differing plausible historic vessels. While this is an undesired outcome, it is a realistic one, and can provide impetus for the targeted collection of additional information from the site. Missing information on the historical vessels is more problematic. While it can always be hoped that a heretofore-unknown historical source will provide new information on a vessel's cargo or its wrecking location, such hoped-for discoveries are often beyond the control of the investigator.

The mention of potential new data in the discussions above highlights a significant strength of the Bayesian approach. New information can easily be incorporated into the probability assessment. If an obscure historical reference is found, or a buried propeller is discovered, the values for all of the ships in the sample can be updated reflecting this new evidence. Similarly, since the probability calculations for each ship are independent of the sample of other vessels, ships can easily be added or removed from the set of plausible vessels. The reporting of results as relative probabilities, and the potential for these posterior probabilities being updated in light of new evidence, underline the probabilistic nature of the historical identifications that are assigned to archaeological shipwreck sites.

Another potential of the Bayesian system is to use the assessment to build a probabilistic description of the vessel that created a wreck site. Such modeling might be employed when a wreck site is newly discovered and a set of plausible historical vessels is not yet developed, or in cases where none of the plausible historically vessels provide a good fit with the observed archaeological evidence. This description, including the vessel's likely length, tonnage, propulsion system, cargo and location of loss, would enable historians to quickly narrow the range of known vessels that might be associated with the wreck.

While the *Bayes Boats 1* system performed well, there is no question that the estimates could be made more decisive by incorporating additional historical information on vessel construction and design. For example, one might evaluate the number of masts on a wrecked vessel. While greater historical detail would allow for better potential discrimination between plausible alternative vessels, the likelihood of actually being able to make the determination on a broken wreck and the uncertainties introduced by ambiguous data both suggest that there are limits to which highly detailed historical information can usefully be added to the model. Yet, even while there may be data limitations to incorporating new categories of historical information, such additions present no problem at the level of the Bayesian system itself. Additional evaluative steps can be added to the system, and the Bayesian network modules for determining conditional probabilities are similarly accessible to modification and refinement.

6. Conclusions

It has been argued in this paper that archaeological inference can benefit from the use of logical and statistical techniques that can handle the uncertainty and imprecision associated with the data collected at wreck sites. A Bayesian approach is well suited to the uncertain and ambiguous datasets inherent to archaeological research. The archaeological applications of the *Bayes Boats 1* system provided useful insights into both how a Bayesian approach might be operationalized, and into the underlying difficulties that scattered wreck sites pose for historical identification.

A Bayesian approach to the identification of scattered wreck sites allows the relative confidence or certainty of historical identifications to be documented and quantified. It equally highlights those cases where expected historical vessels do not make a good fit with the physical evidence, suggesting that a heretofore unsuspected vessel might be represented. The approach can be used to highlight cases where additional archaeological research is necessary and offers a means to update the probabilities of vessel identifications as new information is obtained. It also provides a method for creating an initial descriptive model of the vessel that produced the wreck site.

In place of 'all or nothing' historical identifications, it is now possible to envision a system where the relative probabilities of identifications (and of the archaeological inferences on which they are based) are made explicit and updated in light of ongoing historical and archaeological research. Such a system would encourage new research and assist in focusing research efforts on those wrecks or identifications that are most critically in doubt.

Acknowledgements

This paper has benefited from the comments of Albertus Voorrips, Robert Reynolds, Joyce Marcus, and three anonymous reviewers. Their suggestions have all materially improved the paper, while any shortcomings remain my own.

References

[1] Annual Report of the United States Life Saving Service, Washington, U.S. Government Printing Office, 1890, pp. 40–45.
[2] R. Aykroyd, J. Haigh, G. Allum, Bayesian methods applied to survey data from archaeological magnetometry, Journal of the American Statistical Association 96 (2001) 64–76.
[3] Board of Lake Underwriters, Rules Relative to the Construction of Lake Sail and Steam Vessels, Buffalo, Matthews & Warren, 1866 [reprinted at Maritime History of the Great Lakes, http://www.hhpl.on.ca/GreatLakes/].
[4] C. Buck, J. Kenworthy, C. Litton, A. Smith, Combining archaeological and radiocarbon information: a Bayesian approach to calibration, Antiquity 65 (1991) 808–821.
[5] C. Buck, S. Sahu, Bayesian models for relative archaeological chronology building, Applied Statistics 49 (2000) 423–440.
[6] J. Christen, C. Litton, A Bayesian approach to wiggle-matching, Journal of Archaeological Science 22 (1995) 719–725.
[7] D. Clarke, Models and paradigms in contemporary archaeology, in: D. Clarke (Ed.), Models in Archaeology, Methuen, London, 1972, pp. 1–60.
[8] P. Dellaportas, Bayesian classification of Neolithic tools, Applied Statistics 47 (1998) 279–297.
[9] C. Desmond, Wooden Ship-Building, Rudder, New York, 1919, reprinted 1984, The Vestal Press, Vestal, NY.
[10] Y. Fan, S. Brooks, Bayesian modeling of prehistoric corbelled domes, The Statistician 49 (2000) 339–354.
[11] J. Giarratano, G. Riley, Expert Systems: Principles and Programming, PWS, Boston, MA, 1989, pp. 182.
[12] C. Howson, P. Urbach, Scientific Reasoning: The Bayesian Approach, Open Court, Chicago, IL, 1993, pp. 354–356.
[13] J. Joyce, Bayes Theorem, in: E. Zalta (Ed.), The Stanford Encyclopedia of Philosophy (Fall 2003 Edition), available at: http://plato.Stanford.edu/archives/fall2003/entries/bayes-theorem/.
[14] C. Kadie, D. Hovel, E. Horvitz, MSBNx: A Component-Centric Toolkit for Modeling and Inference with Bayesian Networks, Microsoft Research Technical Report MSR-TR-2001-67, 2001.
[15] C.P. Labadie, Submerged Cultural Resources Study: Pictured Rocks National Lakeshore, Southwest Cultural Resources Center Professional Papers 22, National Park Service, Santa Fe, NM, 1989.
[16] I. Lakatos, in: J. Worrall, G. Currie (Eds.), Philosophical Papers, Cambridge University Press, Cambridge, 1978.
[17] T. Leonard, J. Hsu, Bayesian Methods: An Analysis for Statisticians and Interdisciplinary Researchers, Cambridge University Press, Cambridge, 1999, pp. 96–97.
[18] S.H. Loving, Estimating the age of stone artifacts using probabilities, in: H. Kamermans, K. Fennema (Eds.), Interfacing the Past: Computer Applications and Quantitative Methods in Archaeology CAA 95, Analecta Praehistorica Leidensia no. 28, University of Leiden, Leiden, 1996, pp. 251–261.
[19] J. Lyman, Register tonnage and its measurement, American Neptune V (1945) 223–234, 311–325.
[20] J. Mansfield, History of the Great Lakes, J.H. Beers, Chicago, IL, 1899, pp. 407–424.
[21] K. Muckelroy, A systematic approach to the investigation of scattered wreck-sites, International Journal of Nautical Archaeology 4 (1975) 173–190.
[22] W. Oleszewski, True Tales of Ghosts & Gales, Avery Color Studios, Gwinn, MI, 2003.
[23] J. O'Shea, The archaeology of scattered wreck-sites: formation processes and shallow water archaeology in western Lake Huron, International Journal of Nautical Archaeology 31 (2002) 211–227.
[24] J. O'Shea, Mapping Lake Huron's Au Sable Shoreline and Near Shore Shipwrecks, Report prepared for the Coastal Management Program, Michigan Department of Environmental Quality, Museum of Anthropology, University of Michigan, Ann Arbor, 2002.
[25] M. Salmon, Philosophy and Archaeology, Academic Press, New York, 1982, pp. 49–56.
[26] M. Thompson, Graveyard of the Lakes, Wayne University Press, Detroit, IL, 2000.
[27] J. Timmermann, Sunken Chicora Believed Found, Holland Sentinel, September 21, 2001.
[28] V. van Heest, Searching for a steamer, Michigan History 87 (2003) 10–17.
[29] I. Ward, P. Larcombe, P. Veth, A new process-based model for wreck site formation, Journal of Archaeological Science 26 (1999) 561–570.
[30] A. Wylie, Thinking From Things: Essays in the Philosophy of Archaeology, University of California Press, Berkeley, CA, 2002, pp. 18–19.