https://doi.org/10.1038/s40494-025-01954-x

Prioritizing motion pictures for archival preservation using a decision-aiding Bayesian network

Bruce G. Marcot[1] ✉ **& Mark Ormsby**[2]

The National Archives and Records Administration (NARA) preserves and provides U.S. federal government records of historic and cultural value, including approximately 200,000 reels of 35 mm black-and-white motion picture film. Of immediate concern is allocating finite resources over a 5-year program to preserve the most at-risk, frequently used, and valuable films. To aid this need, we developed a decision-modeling approach based on film deterioration and acidity conditions, known or anticipated access requests, archival value, and other attributes. We used a Bayesian network parameterized with NARA's holdings inventory, a database compiled through laboratory inspection of individual films, and expert judgement of NARA archival and preservation staff. The model aids identifying film series, elements, and reels with high risk ratings and high priority levels, and priorities and risks given film series and element categories, and provides ancillary information for further prioritization decisions and actions, and can be incrementally improved with new data.

The National Archives and Records Administration (NARA) preserves and provides access to billions of permanently valuable records produced by all branches of the U.S. Federal government, including over 500,000 motion picture film reels[1]. Preservation of these items is challenging due to the wide range of media formats and the obsolescence of playback and duplication equipment. In addition, for much of the 20th century, motion picture films were manufactured using a cellulose acetate (CA) base. In poor storage environments, this polymer deteriorates, losing plasticizer and generating acetic acid to produce the odor characteristic of the "vinegar syndrome"[2,3]. As the film becomes more acidic, the reaction reaches an autocatalytic point where the deterioration proceeds more rapidly, and the film eventually becomes unusable. Prior to their transfer to NARA, some films may have been exposed to hot, humid environments for extended periods, leading to the onset of vinegar syndrome[4].

Given the scale of NARA's holdings, the most practical and economical method for preserving CA films is to house them in cold storage, which significantly reduces the rate of acidic deterioration and other degradation reactions[5–7]. In addition, for decades motion pictures have been copied from CA to film stock manufactured on a more stable polyester base[8]. More recently, select films have been digitized and made available online, but this approach is only feasible for a fraction of NARA's motion pictures because of the lack of network infrastructure to support the massive media files.

As detailed below, duplication is driven in part by researcher requests for films for which no suitable backup copy is available. Additionally, groups of films considered particularly valuable, heavily used, seriously deteriorated, or a combination of these factors, are selected for duplication projects. The number of films that can be copied annually is limited to a small portion of the total holdings due to constraints on the available staff, equipment, and funding. Selecting these films for duplication is challenging because of the number of variables involved, as described further below. While NARA utilizes a variety of approaches for evaluating risks to its film collections, these methods make limited use of information available from various sources, in part due to the difficulty of integrating and interpreting this complex mix of data in a decision-advisory context[9,10].

A recent initiative provided an opportunity to produce a flexible, intuitive, decision-aiding tool to address this issue. NARA is in the process of a 5-year project to use its remaining supply of blank 35 mm motion picture stock for film-to-film duplication, including a small fraction of its approximately 200,000 black and white (B&W) holdings. To support this work, we developed a Bayesian network (BN) model to suggest preservation priorities among these films based on decision criteria weighing their physical condition, frequency of use, archival value, and other factors. This approach builds on methods and models of managing information risk[11] that have been used in digital preservation among other institutions, including The National Archives of the U.K[12,13]. and the National Library of New Zealand[14], as well as the DRAMBORA Toolkit ((https://www.repositoryaudit.eu/) for auditing digital preservation repositories. A BN model was selected for this project to tailor the approach to the duplication

[1]USDA Forest Service, Pacific Northwest Research Station, Portland, OR, USA. [2]National Archives and Records Administration, College Park, MA, USA.
✉e-mail: bruce.marcot@usda.gov

project while remaining adaptable to other preservation issues and applications.

Our model development process also serves as a framework for documenting and integrating this information and for analyzing the implications of uncertainties and missing data. Unlike a survey, which gives a snapshot of the information at a given time, the model can easily be updated as more data becomes available and the archival collection continues to expand. The intended audience for this decision-aiding tool includes archivists and preservation staff responsible for these film holdings. This approach may also be applicable to the preservation of audio and video media as well as digital files[15–17]. Here, we describe the methods of developing the database and decision-advisory model, provide several example model assessments for potential prioritization, and suggest future refinements for continued application.

Methods

Archival hierarchies and considerations

The model development process utilizes the structure of NARA's holdings, which are organized in an archival hierarchy of record groups (RG), series, and then individual items[18,19]. A typical RG contains records from a large government entity such as a federal agency or bureau. Within the RG, records are categorized in series which reflect a shared origin, function, or use of the records, often according to the creating office's filing system. For example, RG 255 holds the Records of the National Aeronautics and Space Administration. Among the series filed under RG 255 is FR, "Moving Images Relating to Documentary Stock Footage, 1959–1966" (Local Identifier 255-FR), which consists of footage compiled by the Johnson Space Center for use in internal productions and external requests. A particular item in the FR series might consist of one or more reels of film of various element or format types[20]. An element is an individual roll of film defined by qualities such as whether it is image or sound, negative or positive, color or B&W, etc. Considering only 35 mm B&W motion pictures, NARA has approximately 200,000 films distributed among over 2500 potential combinations of RG, series, and element types. The major decision at hand for NARA thus becomes to determine which film series and elements among this complex mixture should be included in the 5-year duplication project.

Within the RG and series structure, NARA utilizes a 3-level hierarchical organization of film duplicates: a *Preservation copy*, the most original or complete, highest quality copy that is handled as little as possible; an *Intermediate copy* that is used to fulfill requests for various study and presentation purposes; and a *Reference copy*, typically of lower resolution, that is suitable for research and study purposes[21]. Preservation and Intermediate copies are held in cold storage and are handled only by NARA staff or approved vendors. Reference copies are provided directly to researchers at various NARA locations. Ideally, each Preservation copy has an Intermediate that can serve as a backup.

When a researcher requests an item for which no suitable Intermediate or Reference copy is available, the Preservation copy is removed from cold storage and sent to the motion picture preservation lab for inspection prior to eventual duplication. The inspection process includes the use of detector strips which rate the extent of acidic deterioration[22] on a scale of 1–3. The autocatalytic point on the scale is 1.5, above which the deterioration proceeds more rapidly. These acidity readings and other measurements collected during inspections over approximately the past 18 years are stored in a database known as Labwork. Although only a fraction of NARA's films pass through the lab, these data provide insight into the condition of the overall collection as well as information on which series have been requested most frequently over this time.

Whereas all the films in this project are considered permanent federal records that warrant continued preservation, they vary in terms of their archival significance and uniqueness[23]. Significance is evaluated based on four factors: intrinsic significance of the physical item, government significance of content, historical research significance of content, and significance of usage rights and interests. Archival uniqueness of a film includes four factors: whether the film is the only film record source of information it contains; whether the same information is available in other NARA holdings of other record types; whether the information is available in other publicly available sources (published or not); and whether the film source provides additional information not found in any other sources, if other sources do exist.

Data collection and cleaning

As discussed above, the Labwork database contains inspection data collected from 2005 to the present. An export of the data was processed using custom R code[24,25] to address inconsistencies in formatting, changes in series names, and other issues. The data were filtered to focus on only 35 mm B&W films, excluding soundtrack-only elements.

Among over 7500 35 mm B&W films in the Labwork database, only about 1.5% of the reels measured over the past 18 years had acidity ratings at or above the autocatalytic level of 1.5. These results show that, overall, NARA's 35 mm B&W film holdings are in good condition in terms of acidity risk, particularly because the reels are kept in cold storage, which slows the progress of deterioration. The data also provide "real world" evidence to support the predictions of accelerated aging studies that are the basis for utilizing cold storage[7,26–29]. While the overall average risk is low, there may be specific subgroups with higher acidity levels. For example, before being shipped to NARA, films from a particular federal agency may have been stored in poor environments that promoted degradation[4]. Also, among the various B&W element types, some appear to be more susceptible to vinegar syndrome. Identifying these subgroups is challenging given the low overall percentage with high acidity and the large number of variable combinations, which can include hundreds of series, each of which may consist of up to 15 B&W elements.

To narrow the scope, we analyzed the Labwork database to determine which series appeared most frequently over the past 18 years. We found that 11 series accounted for about 70% of the film reels that had been inspected by the lab. Therefore, for prioritizing duplication, we focused on these 11 series rather than attempting to identify high-risk subgroups among the several hundred series in the database. This approach utilized the Labwork database not only for information on acidity but also as a proxy for "use." In an archival context, "high use" is a nebulous concept because a record that is requested a few times a year may be considered high use. Also, interest in certain series might spike due to major anniversaries, changes in world events, scholarly trends, and other unpredictable factors. Nevertheless, we interpreted the frequency of use over the past 18 years – as reflected in the Labwork database – as a reasonable metric for predicting which films are most likely to be requested over the next few decades.

Vinegar syndrome is, however, only one aspect to consider in prioritizing films for preservation. Ideally, every Preservation copy has an associated Intermediate so that there is a high-quality backup. For 35 mm B&W films, roughly 27,000 or 60% are unique, meaning that there is only a Preservation copy, as estimated based on data from NARA's Holdings Management System (HMS) inventory. The analysis was performed using custom software written in R to address gaps in the data, inconsistencies in formatting, and other complications.

To reduce the risk of loss, the 5-year duplication project should prioritize film series that have a high percentage of unique items as well as those considered highly significant. Although film archivists supplied ratings of the archival significance of some series based on written criteria[23], these values are obviously subjective and difficult to replicate precisely and apply consistently to series of widely different content. The evaluation is made more challenging by the lack of information at the item level (i.e., individual film reel) for many series, which may consist of thousands of reels.

In summary, the major factors considered for prioritizing 35 mm B&W films for duplication were aimed at identifying subgroups of specific series and elements that:

1: tended to have higher acidity ratings based on the available Labwork data;

2: were more likely to be requested by researchers;

3: had a higher proportion of Preservation copies without backup Intermediates; and

4: were rated as having high archival value based on their research significance and uniqueness.

Our goal was to develop a risk-assessment decision-advisory tool that denotes levels of preservation priority, that is repeatable, rigorous, and based on records data and that incorporates expert knowledge. The tool would express priority levels as probabilities given uncertainties in film condition, archival values, and other factors. We emphasize that model outcomes are not the final decision on priorities for preservation action but provide decision-advisory information that can be considered by the archivists along with other information not represented in the model, such as budgets, availability of laboratory space and materials, and anticipation of future requests for specific film series or content. In a risk evaluation, model outcomes expressed as probability states can help inform final decisions for risk management in the broader context of such other factors.

Bayesian network advisory model

The BN model structure represents logical and causal relationships among variables (called nodes) that are linked by conditional probabilities calculated using Bayes' Theorem[30,31]. For our purposes, BNs have several key advantages over more traditional risk-analysis models such as decision trees, including being able to transparently combine empirical data with expert judgment[32]; to explicitly express the role and propagation of uncertainty[33]; to deal with missing data without summarily excluding entire sets of data[34]; and to associate risk factors and action priorities in graded levels with probabilities[35]. Unlike other data-analysis modeling structures such as neural networks that mask their pattern-recognition computations, BNs also can show explicitly how each attribute contributes to overall risk and priority levels. Further, BNs can be run both forwards by specifying values of input variables and backwards by specifying potential outcomes or other intermediate variables and determining likely input conditions that result in those outcomes; we found this flexibility to be of great value for helping validate the credibility of the final model and for evaluating the implications of risk and priority levels with archival and preservation experts.

We constructed the BN model using commercially-available software (Netica v. 6.09, Norsys Inc.[36]). We first developed the logical and causal structure of the BN as influence diagrams[37]. These structures integrate known relationships of film conditions using the empirical information summarized from the databases as described above. In addition, the influence diagrams incorporate expert knowledge and experience through judgments of risk levels based on interviews with NARA staff working in accessioning, collections management, and film preservation and conservation. We initially created numerous draft versions of influence diagrams with all these factors and had NARA staff review them for logic and representation of their risk evaluation approaches. We amended a final diagram to best conform to their suggestions following general procedures for developing model structures from expert knowledge[38].

The BN model consists of 6 variable group network submodels with varying numbers of variables shown in the network as nodes, with each node represented by discrete or continuous-value states (Table 1). We used our film record data to identify state values and ranges for continuous value variables by applying the program's discretization algorithms and used our archivist interviews to identify the discrete states for the series archival values variables and the risk assessment variables (see Supplemental Information). We then parameterized the conditional probability tables in the BN model with the film record data by using the expectation maximization algorithm, which consists of a stepwise convergent log-likelihood function that assigns probability values in the network that best conform to known data outcome values[39].

Once we fully parameterized the network with unconditional and conditional probability values, we ran the model forward to determine risk and priority levels for specified film series and elements, and backward to identify series and elements with specified risk and priority levels. We assessed the calibration accuracy of the model by comparing the predictions of the series and element descriptors (acidity level index, used based on Labwork percentage, and percent unique Preservation copies) to the known values of these same variables in the database, by film series and elements. We also conducted a sensitivity analysis on the final model to determine the degree to which priority levels are sensitive to input and intermediate variables.

Results

Here we present results of our compilation of data on film series and elements, and on the structure and results of the Bayesian network decision-advisory model on potential priority ratings of film series and elements.

Model structure

The model was structured so that specifying film series and/or element will, in turn, denote their attributes that summarize into priority ratings. The final BN model uses six categories of variables, noted here as variable group submodels (Fig. 1). The *Series and Element submodel* consists of 11 series (those that made up 70% of the Labwork entries) and each of the 15 possible B&W element types. The combination of series and element provides unique attributes for the rest of the model, as indicated by the arrows that lead to the sets of record attributes (archival value, and record group and element descriptors) and outcomes (priority and risk ratings, and ancillary information) used in, and resulting from, the analysis.

Next, the *Series Archival Values submodel* consists of archival significance and archival uniqueness ratings of each series, which are scored on categorical scales of 1 to 3 of increasing significance and uniqueness. These components are then summed into an overall archival value score that essentially provides equal weight to archival significance and archival uniqueness, following the risk perspective of the archivists. The *Series and Element Descriptors submodel* consists of three components: the percent of unique Preservation copies in NARA's HMS inventory; an estimate of the

Table 1 | Submodel structure of the Bayesian network (see Fig. 1) for prioritizing 35 mm black-and-white archival film series and elements for preservation action

Network Submodel	Description	Information Source
Series and Element	Lists the specific categories of the film series and elements	Database compiled from HMS.
Series and Element Descriptors	Describes uniqueness, use, and acidic deterioration of films	Database compiled from Labwork and HMS.
Series Archival Values	Describes the archival significance and uniqueness of films, and their composite archival value rating	Significance and uniqueness are assigned by archival staff based on criteria in Reference [7]; composite archival value rating is a sum of these
Priority and Risk Ratings	Summarizes film conditions and archival values into risk and priority ratings	Based on input from NARA archivists
Series and Element Ancillary Information	Additional descriptions of film Series and Element conditions, Labwork, and numbers of available copies	Database compiled from Labwork and HMS.
Case Number	Tracks the specific case number in the case file database	Database compiled from HMS

HMS Holdings Management System of NARA.

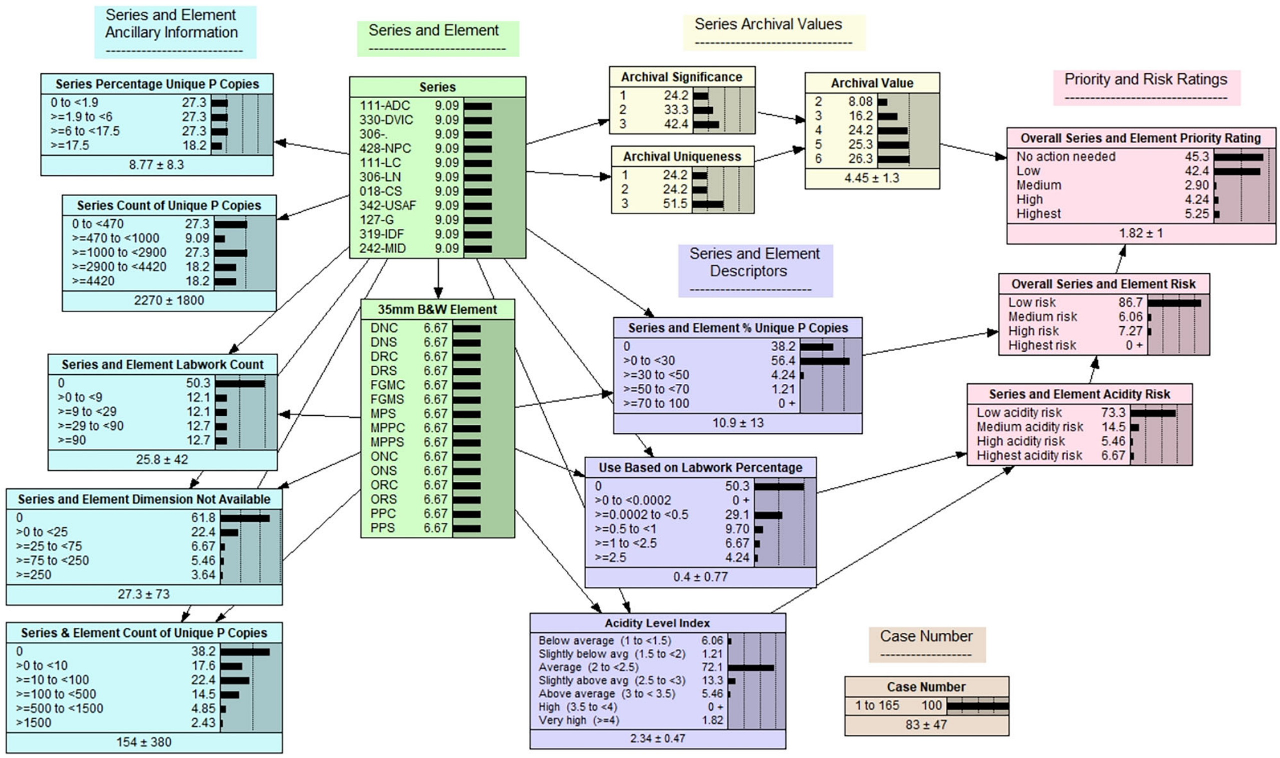


Fig. 1 | Bayesian network model providing overall priority ratings for potential preservation action for 35 mm B&W film records for NARA. Values by the black bars denote probabilities of each state condition and sum to 100 percent for each node. Values below the continuous-value nodes are expected values +/− 1 standard deviation given the state ranges and associated probabilities. (See Supplementary Information for examples of specific model runs.)

use of each film series and element based on the percent of the total Preservation copy entries in Labwork across all series and elements; and an acidity level index of each film series and element, based on the shape of the acidity distribution curve as derived from Labwork data. The acidity index was assigned based on inspection of acidity data histograms and covers a range 1–4 in intervals of 0.5, not to be confused with the 1–3 scale of the detector strips. The average acidity risk was set at 2.0, and higher values represent greater risk. Series and element combinations lacking sufficient data to evaluate the acidity level were assigned the default average value of 2.0. The *Priority and Risk Ratings submodel* provides the assessments of overall potential priority ratings, as informed by archival value and by overall series and element risk, and as denoted on a five-class ordinal scale ranging from no action needed to highest priority rating. In total, the Bayesian network overall risk and priority ratings model consists of 17 nodes, 25 links, and 6,345 conditional and unconditional probability values.

The model also provides additional profiles in the *Series and Element Ancillary Information submodel*. These are not directly linked to the priority and risk ratings but provide additional information that may be useful for evaluating model outputs and guiding preservation decisions. They include five ancillary information categories: Series percentage unique Preservation copies, which is the percentage of unique Preservation copies of all element types specified for a given series; Series count of unique Preservation copies, which denotes the sum of unique Preservation copies of all element types for a specified series; Series and Element Labwork count, which is the number of entries of Preservation copies in Labwork documentation; Series and Element dimension not available, which is the number of entries of specified series and element types where the film gauge dimension (generally 16 mm or 35 mm) is not listed in the HMS inventory; and Series and Element count of unique Preservation copies, which is the number of unique Preservation copies (that is, lacking Intermediate copies) in the HMS inventory among all element types for a specified series (and, since dimension is unknown for some HMS inventory entries, the total may include some 16 mm film stock records). In addition, the *Case Number submodel* essentially displays the specific record group series and element type entry in the case file used to parameterize and calibrate the model, as the model is run specifically from that file (but can also be run independently).

Results of running the model

Here we present a few examples of using the model to provide insights on the kinds of results and inferences that can be made regarding film archival priority, risk levels, and conditions. Table 2 presents two examples of each way to run the model (also see Supplementary Information for model images of each example run). In running the model forward, one can specify the film record series and/or element, and determine the resulting potential archival priority, film risk levels, and associated film attributes that can help explain the results. For instance, example 1 starts by selecting series 330-DVIC and element ONS (original negative silent, see ref. 20). These choices result in the highest archival priority and high overall risk, in large part because this combination has the highest ratings for acidity risk and for archival significance and uniqueness. Examining the Labwork count node shows that the acidity level assignment is based on a large number of samples, which bolsters confidence in this assessment. Example 2 (Table 2) demonstrates that specifying film series 111-ADC and element ONS results in identifying those film records as having low acidity and low overall risk. For this combination the Labwork count is small (< 9), which may raise concern because there was limited data on which to base the acidity index. Both the percentage of unique P copies and the use levels are in the mid-to-low range, however, which supports the overall low priority rating for these films for the 5-year duplication project.

In running the model backward, one can specify the priority and risk rating levels to determine which film series and records are implicated, as well as other attributes of those film records. Example 3 (Table 2) shows that specific film series (330-DVIC) and elements (MPPC) can be identified as those records having the highest overall Series and Element risk level, very high acidity level as evidenced by highest Labwork percentage. Example 4, on the other hand, determines that specifying just the penultimate level of

Table 2 | Examples of running the Bayesian network model for prioritizing film records for preservation action (see Fig. 1) with explanations of interpreting results

Node and State Conditions Specified \a	Node Outcomes of Key Results and Interpretation
Example of running the model forward	
Example 1: Series: "330-DVIC" and Element: "ONS - Original Negative, Silent"	Key results: • Overall Series and Element Priority Rating: 100% "Highest" • Overall Series and Element Risk: 100% "High risk" • Series and Element Acidity Risk: 100% "Highest acidity risk" • Acidity Level Index: 100% "Above average" • Use Based on Labwork Percentage: 100% "> = 2.5" • Acidity Level Index: 100% "Above average" • Archival value: 100% "6" Interpretation: These records are seen as highest priority and high overall risk, because acidity risk is highest due to highest use, acidity level index is above average, and archival value of significance and uniqueness is highest.
Example 2: Series: "111-ADC" and Element: "ONS - Original Negative, Silent"	Key results: • Overall Series and Element Priority Rating: 67% "No action needed" and 33% "Low" • Overall Series and Element Risk: 100% "Low risk" • Series and Element Acidity Risk: 100% "Low acidity risk" • Acidity Level Index: 100% "Average" • Series and Element % Unique P Copies: 100% ">0 to <30" • Use Based on Labwork Percentage: 100% "> = 0.0002 to <0.5" • Acidity Level Index: 100% "Average" Interpretation: These Records are seen as low priority and low risk, because of low uniqueness, low use, and only average acidity level.
Example of running the model backward	
Example 3: Overall Series and Element Risk: "Highest risk" Acidity Level Index: "Very high" Use Based on Labwork Percentage: "> = 2.5" Series and Element % Unique P Copies: "> = 70 to 100"	Key results: • Series: 100% "330-DVIC" • Element: 100% "MPPC - Master Projection Print, Composite" • Series and Element Labwork Count: 100% "> = 90" • Series and Element Dimension Not Available: 100% "0" Interpretation: Identification of the series and element that conform to the specified attribute of highest risk is correct and trustworthy, and the associated records are of the appropriate type for consideration, because: the ancillary information shows the highest level of Labwork (highest empirical data basis) which lends greatest credence to the specified very high acidity rating; and all records are of the target 35 mm dimension (because dimension not available is 0).
Example 4: Series and Element & Unique P Copies: "> = 50 to <70"	Key results: • Series: 50% "306-LN", 50% "018-CS" • Element: 50% "FGMS - Fine Grain Master, Silent", 50% "MPPC - Master Projection Print, Composite" • Overall Series and Element Priority Rating: 56% "High", 28% "Medium, 17% "Low" • Overall Series and Element Risk: 100% "Medium risk" • Series and Element Acidity Risk: 100% "Low acidity risk" Interpretation: These identified series and elements carry medium Overall Series and Element Risk, and mostly medium to high priority, as evidenced by below-average acidity level and relatively high Labwork count.

Running the model forward refers to specifying series and/or element, and exploring results of priority and risk ratings and other factors; *running the model backward* refers to specifying any condition other than series and element, and exploring which series and/or elements are implicated and other associated factors. Percentages from the model are rounded to nearest whole units. For a list of element abbreviations see ref. 20. (See Supplementary Information for model images of each example run).
\a Series and Element codes used in this table:
Series 330-DVIC: Record Group 330 = Records of the Office of the Secretary of Defense; DVIC = Motion Picture Films and Video Recordings on Five Decades of U.S. Military Activities Around the World, ca. 1950 to ca. 2000.
Series 111-ADC: Record Group 111 = Records of the Office of the Chief Signal Officer; ADC = Moving Images Relating to Military Activities.
Series 306-LN: Record Group 306 = Records of the U.S. Information Agency; LN = Motion Picture Films from the "Liberty News" Program Series, 1953-1967.
Series 018-CS: Record Group 018 = Records of the Army and Air Forces; CS = Motion Picture Films from the "Combat Subjects" Program Series, ca. 1939 to ca. 1945.
Element ONS: Original Negative Silent (film).
Element MPPC: Master Projection Print, Composite.
Element FGMS: Finegrain Master, Composite.

Series and Element Percent Unique Preservation copies can reveal that two series (306-LN and 018-CS) and two elements (FGMS and MPPC) pertain, and that they carry overall high to medium risk levels but low acidity risk.

In this way, NARA staff can apply their expert knowledge and intuition to run the model and explore a wide variety of queries to help reveal film series and element types of various priority and risk levels and the empirical basis for such determinations. The examples above produced relatively clear outcomes, but when dealing with missing data or potentially contrary conditions, results may be expressed more as a spread of probabilities of outcome states. In such cases, the spread may be interpreted as useful information on uncertainty, and the model can be used to determine what attributes of the film records are unknown or potentially deserving of additional attention. For example, if the Series and Element Labwork Count were relatively low – indicating that the Acidity Level Index was assigned based on limited data – then additional lab inspections might be ordered to gather supplemental acidity readings and more carefully assess the risk.

Model evaluation

We reviewed potential network structures during model development with four archivists and motion picture preservation specialists (see Acknowledgments). Their reviews were conducted independently, by email and in person, and were sought after we completed the fully structured and parameterized draft model. We asked them specifically if there were any key parameters in the model for input, analysis, and output, missing or misidentified, and if the state categories in the model were usefully depicted for the priority and risk ratings variables. Their responses largely supported the draft model as presented, with minor suggestions to improve clarity of presentation. We then also sought their review of the final model structure and its operation and performance as a test of "face validity"[40], that is, if the model made logical sense and provided results and information that conformed to the experts' experience. As we incrementally amended the model structure and parameters from their input, the final model passed the expert evaluation reviews.

Table 3 | Sensitivity analysis of the Bayesian network model for prioritizing film records for preservation action (Fig. 1)

Node	Variance Reduction	Percent
Overall Element Risk	0.7961	73
Element Acidity Risk	0.7542	69.1
Acidity Level Index	0.5928	54.3
Archival Value	0.4466	40.9
Series	0.4024	36.9
Use Based on Labwork Percentage	0.2878	26.4
Series and Element Labwork Count[a]	0.2779	25.5
Archival Uniqueness	0.2719	24.9
Archival Significance	0.1875	17.2
Series & Element Count of Unique P Copies[a]	0.1436	13.2
Series Count of Unique P Copies[a]	0.1400	12.8
Element	0.1063	9.75
Series and Element Percent Unique P Copies	0.0849	7.78
Series Percentage Unique P Copies[a]	0.0667	6.11
Series and Element Dimension Not Available[a]	0.0454	4.16

Results denote the relative sensitivity of the Overall Priority Rating outcome node to all other nodes in the network. (P Copies = Preservation Copies). Variance reduction values denote the degree to which the Overall Priority Rating outcome node is sensitive to each specified node in the model (see Marcot 2012 for formula used).
[a]Nodes representing indirect effects of ancillary information; see text for description.

Another consideration for model authenticity regards calibration accuracy, that is, how well a model conforms to a set of data used to develop its structure and/or parameters. Regarding our model, achieving high calibration accuracy was a trivial outcome, with 0% error, in the sense that the available data provided but one case example for each combination of film record series and element with no multi-case stochastic variation. As new data become available, which is indeed the situation with NARA's HMS and Labwork databases, the model can be retested for calibration accuracy and retrained with the additional data, building upon the previous model version.

We also conducted a model sensitivity analysis to determine the degree to which the final outcome node in the model, Overall Series and Element Priority Rating, was sensitive to each other node. Results (Table 3) suggest that key sensitivities pertain to nodes nearest the outcome node (which seems logical but is not necessarily given and depends on the probability structure of the model), namely: Overall Element Risk, Element Acidity Risk, Acidity Level Index, Archival Value, and series, for the top 5 nodes (with variance reduction > 0.4 and sensitivity percent > 35; Table 3). However, the analysis showed that the outcome node had some degree of sensitivity to all nodes in the model, suggesting that none was superfluous.

Discussion

A major value of this decision-advisory modeling tool is being able to ask a variety of questions and to conduct diverse queries, such as solving forward from film series and elements, and backwards from priority ratings, risk levels, or any intermediate outcomes. Essentially, this provides a means of running multiple scenarios of record conditions, risk levels, and priorities. Previously, the raw data were available in various sources, but they were not integrated in a way that allowed NARA staff to exploit them for deeper insight.

Our approach of structuring and parameterizing a Bayesian network model to evaluate priority ratings and risk levels of records is flexible enough to work with relatively small databases. Whereas some collections of raw data included thousands of entries, the model itself was parameterized from a spreadsheet with 165 cases that summarize larger databases. A main advantage of the decision-advisory network model is to express uncertainties of outcomes, when available data cannot provide perfect causal relationships. The type and degree of uncertainty then is a key factor in risk-management decision-making for preservation action, such as when weighing the model outcomes of probabilities of the various levels of element risk and potential decision priorities.

Decision-advisory systems, including use of BN models[41,42], are often developed to advise and guide environmental conservation and classification actions such as with aerial imagery[43]. More directly relevant to the current work is that of Barons et al.[44] that provided an interactive tool in the National Archives of the U.K., as an interactive, decision-support system, based on a BN model, that quantifies potential risks associated with their preservation of digital material. Their decision-support tool is the Digital Archiving Graphical Risk Assessment Model (DiAGRAM)[45], which was developed independently of our work.

The next step in implementation is providing the model to archivists and preservation lab staff personnel working on the 5-year duplication project. We also suggest tracking the degree to which this decision-advisory tool can provide specific cost savings by optimizing use and allocations of finite resources, lab space, and available personnel through the prioritization process.

Next steps can also include incrementally updating the model parameters, and potentially revising the network structure, from newly-entered database cases. The R code was written to automate as much of this process as possible so that new data can be easily incorporated. Updating already occurred during model development. When six months of new Labwork data were imported, the values for the Acidity Level Index and Use Based on Labwork Percentage for some series and element combinations were revised. This result was somewhat surprising given that the proportion of new entries was small compared to the 18 years-worth of data already incorporated into the model. The intuitive, graphical structure of the BN enables archival staff to explore the implications of these changes. This, indeed, is the 'spirit' of Bayesian statistics by using updated prior information to revise posterior predictions. New data could also be used to conduct specific validation tests of the model.

Our BN decision-advisory model integrates a wide variety of complex and often incomplete data into an intuitive, visual tool to assist NARA in prioritizing films for duplication. The graphical interface and ability to run the model forward and backward allow archivists and preservation staff to analyze the data, explore scenarios, and gain insights into the relationships between multiple variables. This approach also serves as a framework for adapting to other media types as a decision-advisory tool.

Data availability

The Bayesian network model and database are available at https://doi.org/10.59381/kshcfvxqsn.

Code availability

The R code mentioned that was developed specifically to structure the NARA inventory database is available from the corresponding author upon request.

Received: 12 September 2024; Accepted: 14 July 2025;
Published online: 29 July 2025

References

1. National archives by the numbers. https://www.archives.gov/about/info/national-archives-by-the-numbers. Accessed 23 October 2024.
2. Allen, N. S., Edge, M., Appleyardt, J. R., Jewitt, T. S. & Rorie, C. V. Degradation of historic cellulose triacetate cinematographic film: influence of various film parameters and prediction of archival life. *J. Photo. Sci.* **36**, 194–198 (1988).
3. Adelstein, P. Z., Reilly, J. M., Nishimura, D. W. & Erbland, C. J. Stability of cellulose ester base photographic film. *part III Meas. film. Degrad. Soc. Motion Pic Telev. Eng. J.* **104**, 281–291 (1995).

4. Comptroller General's report to Congress, Valuable government-owned motion picture films are rapidly deteriorating. General Accounting Office, Washington, D.C. June 19, 1979;LCD 78-113.
5. McCormick-Goodhart, M. H. The allowable temperature and relative humidity range for the safe use and storage of photographic materials. *J. Soc. Archivists* **17**, 7–21 (1996).
6. Reilly, J. M. IPI storage guide for acetate film. Rochester: Image Permanence Institute. 1993.
7. Richardson, E. J., Cummings, M. & Bigourdan, J.-L. Context, development, and intent: an introduction to the IPI preservation metrics. *Heritage* **6**, 4202–4213 (2023).
8. National Film Preservation Foundation (US). The film preservation guide: the basics for archives, libraries, and museums. National Film Preservation Foundation. 2004.
9. Thibodeau, K. Breaking down the invisible wall to enrich archival science and practice. In proceedings of IEEE International Conference on Big Data. 2016.
10. Richards, J. & Brimblecombe, P. The transfer of heritage modelling from research to practice. *Herit. Sci.* **10**, 17 (2022).
11. Saffady, W. Managing information risks: threats, vulnerabilities, and responses. Rowman & Littlefield Publishers, 2020;256 pp.
12. Underdown, D. & Merwood, H. Quantifying digital preservation risks using statistics. *IQ: RIM Q.* **36**, 42–43 (2020). https://search.informit.org/doi/abs/10.3316/informit.475513445135716.
13. McHugh, A. A model for digital preservation repository risk relationships. Proceedings of World Library and Information Congress: 78th IFLA General Conference and Assembly. Helsinki, Finland. 2012; http://eprints.gla.ac.uk/65420.
14. De Vorsey, K. & McKinney, P. Digital preservation in capable hands: taking control of risk assessment at the National Library of New Zealand. *Inf. Stand. Q.* **22**, 41–44 (2010).
15. Lewis, D. R. Making sound decisions: institutional responses to the crisis in audio preservation. *Archival Iss* **40**, 23–44 (2020).
16. Kim, J. et al. Audiovisual quality control and preservation case studies from libraries, archives, and museums. *Intern. Assoc. Sound Audiovis. Arch. ((IASA)) J.* **51**, 23–40 (2021).
17. Johnston, L. Implementing a framework for digital preservation risk assessment and mitigation at the US National Archives. *J. Digit Media Mgmnt* **8**, 351–360 (2020).
18. Record group concept. National Archives. https://www.archives.gov/research/guide-fed-records/index-numeric/concept.html. Accessed 23 October 2024.
19. Trace, C. B. Maintaining records in context: a historical exploration of the theory and practice of archival classification and arrangement. *Am. Arch.* **83**, 91–127 (2020).
20. National Archives and Records Administration (NARA) motion picture technical processing guidance. https://www.archives.gov/files/preservation/formats/pdf/nara-motion-picture-technical-processing-guidance.pdf. Accessed 23 October 2024.
21. Kovac C., Love, J. Audio-visual collection preservation at the NARA. Against the Grain 2015;27:Article 11. https://doi.org/10.7771/2380-176X.7129.
22. Using A-D strips. http://filmcare.org/ad_strips. Accessed 23 October 2024.
23. Supplement 3 – determining the significance of NARA holdings. https://www.archives.gov/files/foia/1571-supplement-3-significance-nara-holdings.pdf. Accessed 23 October 2024.
24. The R Project for statistical computing. https://www.R-project.org. Accessed 23 October 2024.
25. Grolemund, G., Wickham, H. R for data science. Sebastopol, CA: O'Reilly Media. 2017.
26. Adelstein, P. Z., Reilly, J. M. & Emmings, F. G. Stability of photographic film: Part VI—Long-term aging studies. *SMPTE J.* **111**, 136–143 (2002).
27. Bigourdan J.-L.Stability of acetate film base: accelerated-aging data revisited. Archiving Conf. Vol. 2. Soc Imaging Science and Technology. 2005.
28. Knight, B. Lack of evidence for an autocatalytic point in the degradation of cellulose acetate. *Polym. Degrad. Stab.* **107**, 219–222 (2014).
29. Ahmad, I. R. et al. Are we overestimating the permanence of cellulose triacetate cinematographic films? A mathematical model for the vinegar syndrome. *Polym. Degrad. Stab.* **172**, 109050 (2020).
30. Heckerman, D. A tutorial on learning with Bayesian networks. Springer NY. Innovations in Bayesian networks (Vol. 156). Studies in Computational Intelligence. 2008. p 33-82. https://doi.org/10.1007/978-3-540-85066-3_3.
31. Koski, T., Noble, J. Bayesian networks: an introduction. John Wiley & Sons, West Sussex, U.K. 2011. 368 pp.
32. Constantinou, A. C., Fenton, N. & Neil, M. Integrating expert knowledge with data in Bayesian networks: preserving data-driven expectations when the expert variables remain unobserved. *Expert Syst. Appl.* **56**, 197–208 (2016).
33. Marcot, B. G. Metrics for evaluating performance and uncertainty of Bayesian network models. *Eco Mod.* **230**, 50–62 (2012).
34. Chen, S. H. & Pollino, C. A. Good practice in Bayesian network modelling. *Env Mod. Softw.* **37**, 134–145 (2012).
35. Garvey, M. D., Carnovale, S. & Yeniyurt, S. An analytical framework for supply network risk propagation: a Bayesian network approach. *Eur. J. Op. Res.* **243**, 618–627 (2015).
36. Norsys Software Corp. https://www.norsys.com/. Accessed 23 October 2024.
37. Kjaerulff. U. B., Madsen. A. L. Bayesian networks and influence diagrams: a guide to construction and analysis. Springer, New York. 2007. 318 pp.
38. de Waal A. et al. Construction and evaluation of Bayesian networks with expert-defined latent variables. 2016. P. 774-781 in: 19th International Conference on Information Fusion (FUSION). Curran Associates, Heidelberg, Germany.
39. Do, C. B. & Batzoglou, S. What is the expectation maximization algorithm?. *Nat. Biotech.* **26**, 897–899 (2008).
40. Nevo, B. Face validity revisited. *J. Educ. Msmnt* **22**, 287–293 (1985).
41. Haas, T. C. A Bayesian belief network advisory system for aspen regeneration. *Sci* **37**, 627–654 (1991).
42. Marcot, B. G., Hoff, M. H., Martin, C. D., Jewell, S. D. & Givens, C. E. A decision advisory system for identifying potentially invasive and injurious freshwater fishes. *Mgmnt Bio Invasions* **10**, 200–226 (2019). https://www.proquest.com/scholarly-journals/decision-support-system-identifying-potentially/docview/2285114813/se-2?accountid=12693.
43. Movia A., Beinat A., Sandri T. Land use classification from VHR aerial images using invariant colour components and texture. 2016. Pp. 311-317 in: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol XLI-B7, 2016 XXIII ISPRS Congress, 12–19 July 2016. Prague, Czech Republic.
44. Barons, M. et al. Safeguarding the nation's digital memory: towards a Bayesian model of digital preservation risk. *Arch. Rec.* **42**, 58–78 (2021).
45. DiAGRAM - The Digital Archiving Graphical Risk Assessment Model. https://diagram.nationalarchives.gov.uk/. Accessed 23 October 2024.

Acknowledgements

We thank Christina Austin, Jennifer Herrmann, and Ellen Mulligan of NARA for their helpful reviews of the manuscript and model structure. We thank Christina Austin, Martin Jacobson, and Ellen Mulligan for their guidance and their reviews of the model specifically for archival risk analysis and application, and Steve Graybill of NARA for assistance in exporting HMS and Labwork data. BGM acknowledges support from USDA Forest Service, Pacific Northwest Research Station. Mention of commercial products does not necessarily imply endorsement by U.S. Government. No generative AI tool was used in the analysis or writing of the manuscript

Author contributions

Conceptualization: B.G.M. and M.O.; methodology: B.G.M. and M.O.; database: M.O.; network modeling: B.G.M.; writing -- original draft preparation, review, and editing: B.G.M. and M.O.; funding acquisition: M.O. Both authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s40494-025-01954-x.

Correspondence and requests for materials should be addressed to Bruce G. Marcot.

Supplementary Information

for: " Development of Bayesian network decision-aiding tools to prioritize archival records for preservation."
by: Bruce G. Marcot, Mark Ormsby, Ellen Mulligan, Christina Austin, and Martin Jacobson

Heritage Science

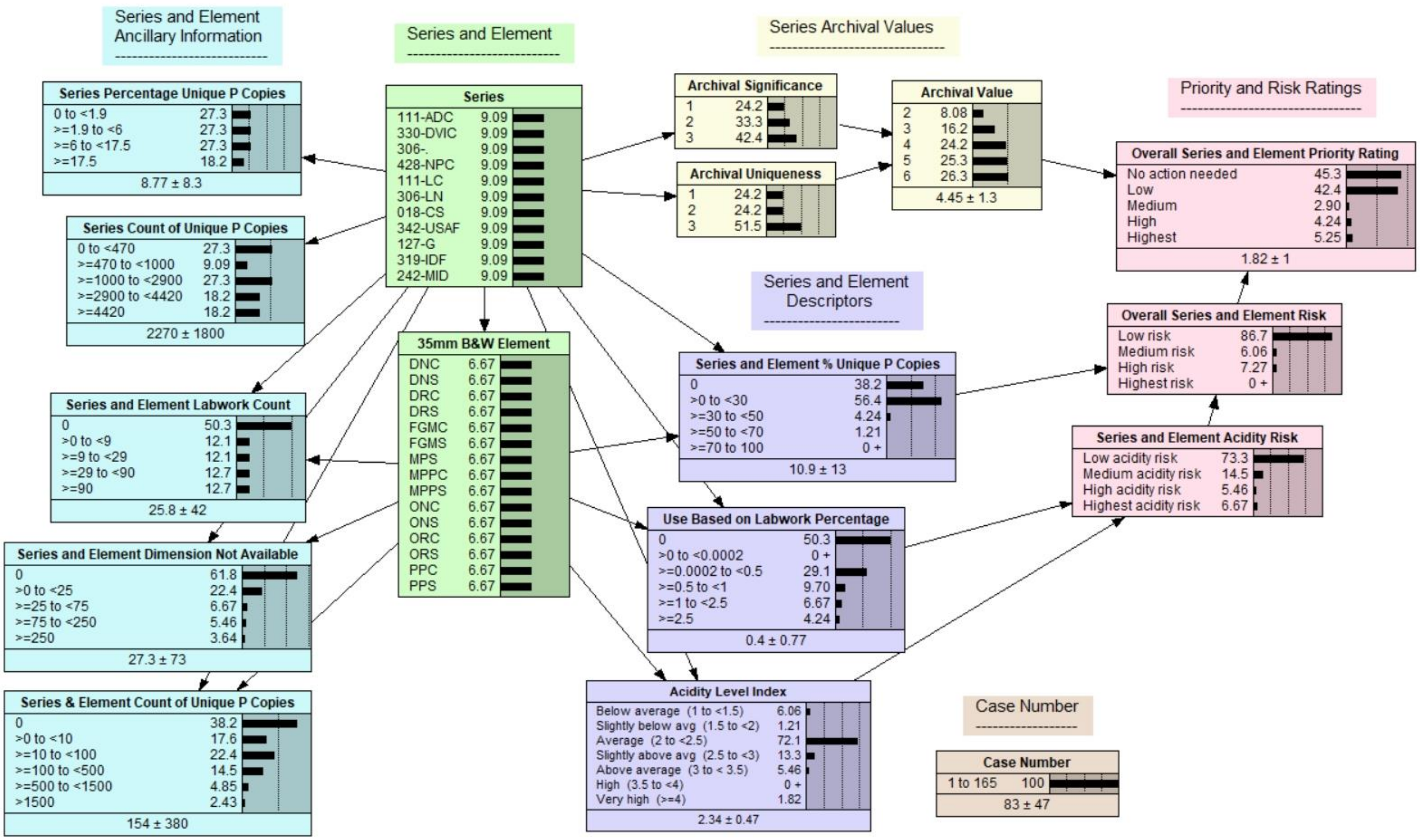
==================

Supplementary Information.

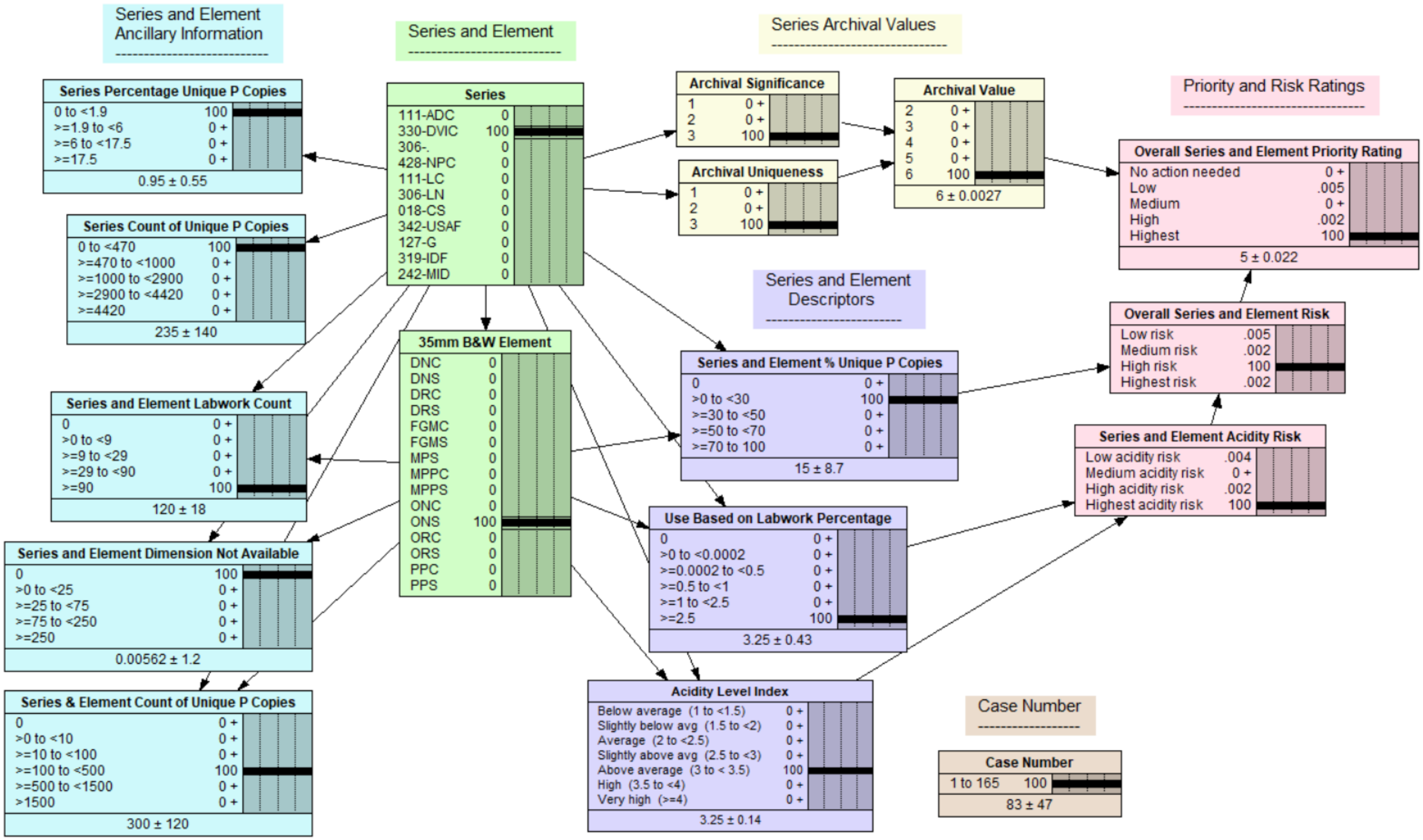
Figures: Four examples of running the Bayesian network model denoting risk levels to prioritize 35mm B&W (black and white) film archival records for preservation. See main text Table 2 for each numbered example run specifications, specific results highlighted, and interpretations of results for archival prioritization. A full list of abbreviations for elements is available in main test Reference [20].

Tables: Definitions of nodes and states used in the Bayesian network model of archival 35mm B&W film priority ratings (see text for explanation) for National Archives and Records Administration (NARA), USA. The case file database mentioned is referenced in the main text.

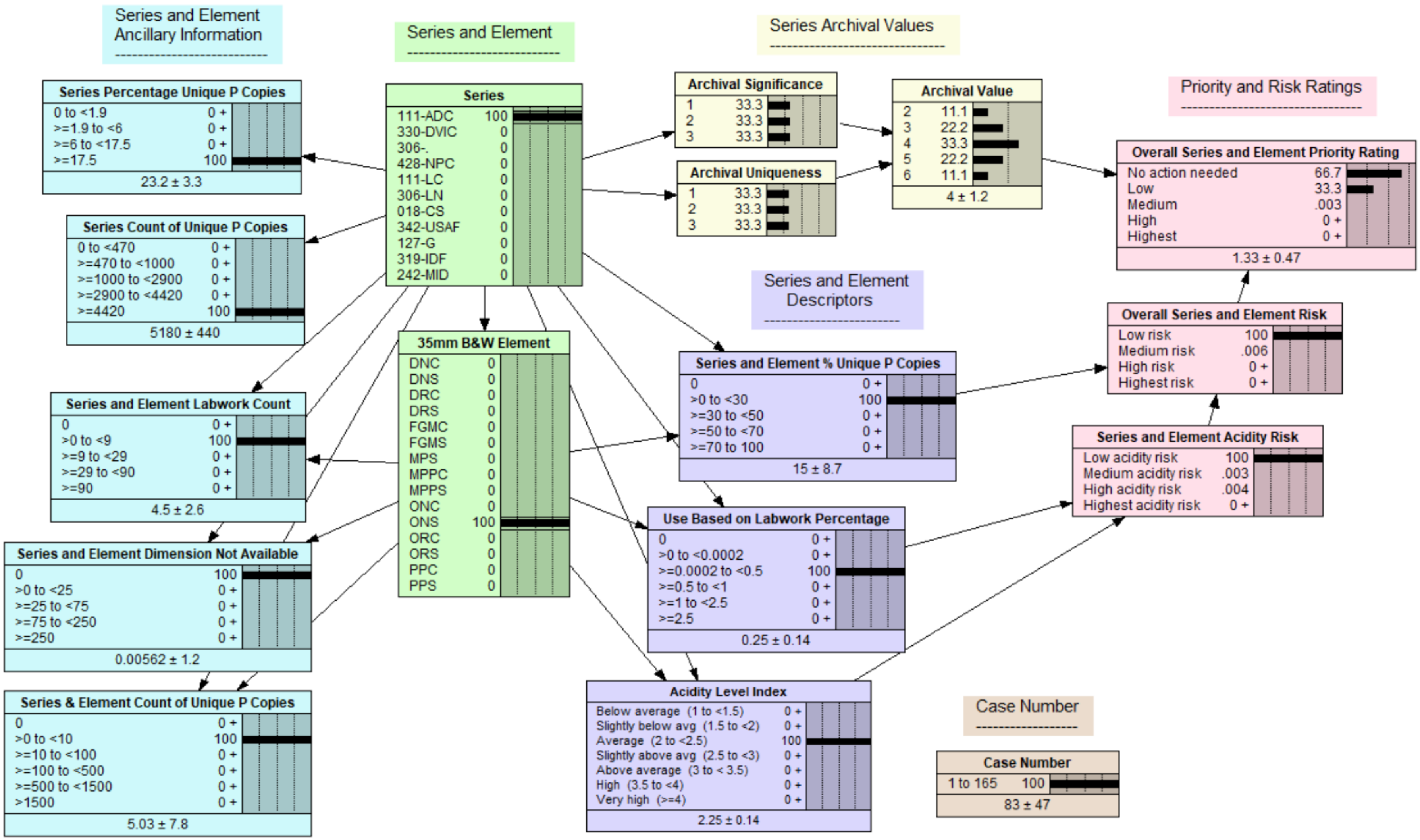
Supplementary Information Figure 1. Initial default state of the model. This figure shows the initial, default state of the model. Values by the black bars denote probabilities of each state condition and sum to 100 percent for each node. For example, in the initial state, all 11 series have equal probability of 9.09%. When the user selects a particular state within a node, the model updates and shows the probabilities in all nodes that match that specified situation.



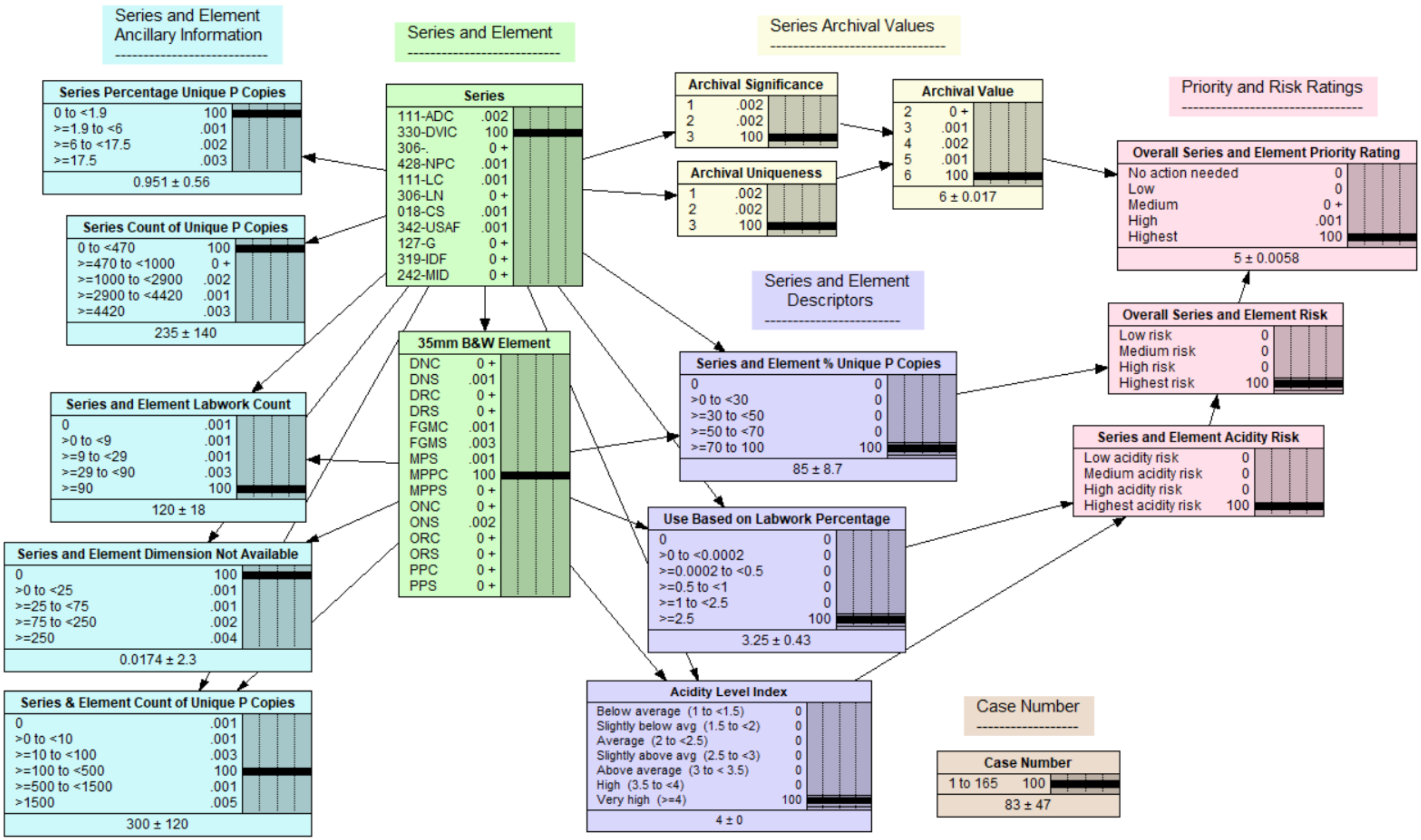
Supplementary Information Figure 2. Example 1: Running the model forward, specifying Series: "330-DVIC" and Element: "ONS". Once the user selects series 330-DVIC and element ONS, those nodes indicate 100% for those conditions and 0% for all others. The model updates the rest of the nodes based on these selections, showing that this combination of series and element has the highest overall priority rating.



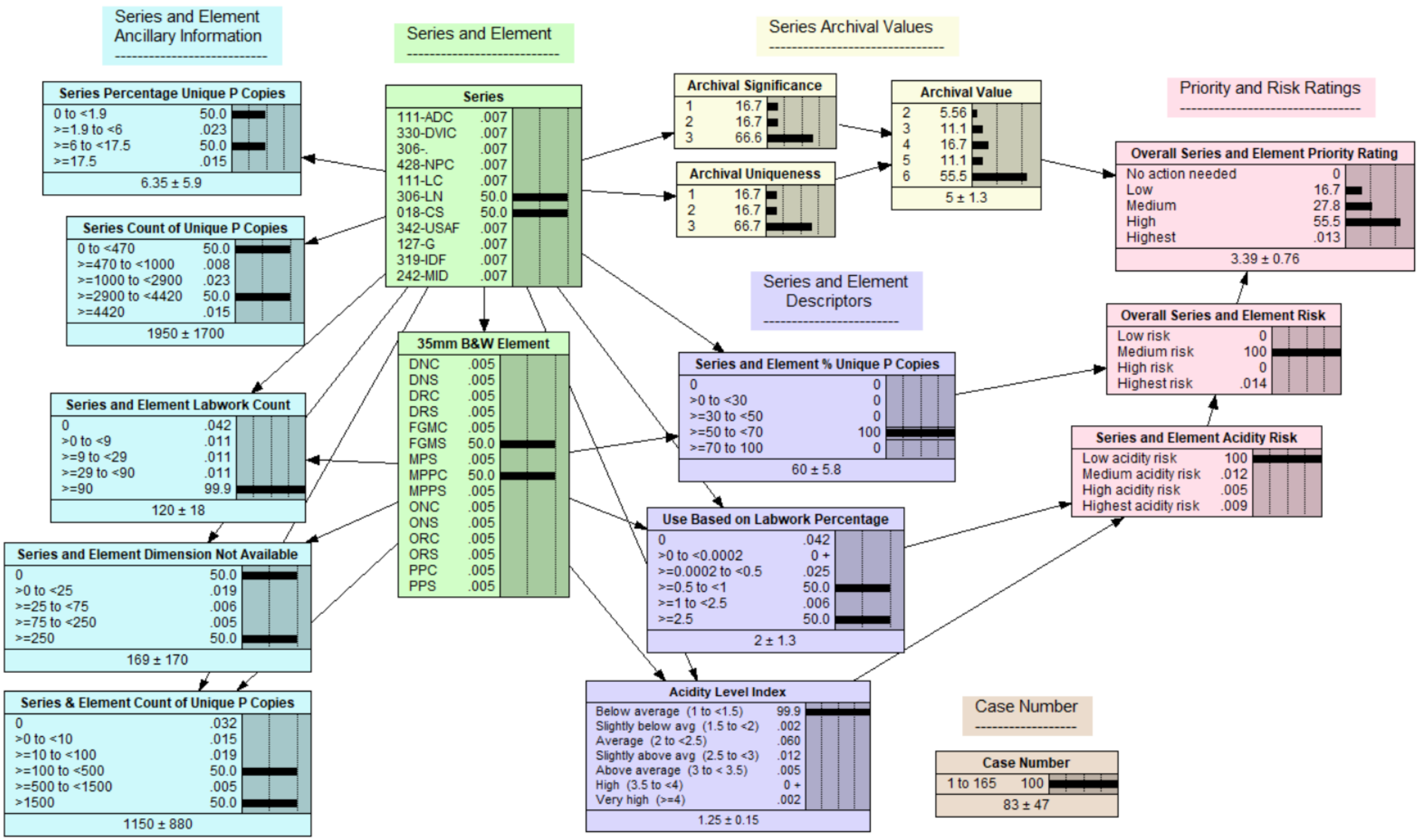
Supplementary Information Figure 3. Example 2: Running the model forward, specifying Series: "111-ADC" and Element: "ONS". Once the user selects series 111-ADC and element ONS those nodes show 100% for those conditions and 0% for all others. The rest of the nodes are updated based on these selections.



Supplementary Information Figure 4. Example 3: Running the model backward, specifying Overall Series and Element Risk: "Highest risk"; Acidity Level Index: "Very high"; Use Based on Labwork Percentage: ">=2.5"; Series and Element % Unique P Copies: ">=70 to 100". Working backward, the user selects the highest option for all nodes in the Priority and Risk Ratings and Series and Element Descriptors submodels. Based on those selections, the model updates and identifies the highest priority as element MPPC in series 330-DVIC.



Supplementary Information Figure 5. Example 4: Running the model backward, specifying Series and Element % Unique P Copies: ">=50 to <70". In this case, the user is interested in series and element combinations that have a large percentage (50 to 70%) of P copies that are unique. Selecting that condition causes the model to update, and it identifies two series and two elements that match this criterion. It also shows that half of the reels are at the highest level of use (>=2.5). The combination of high use and relatively large percentage of unique copies may be a concern for at least some of the reels. However, the acidity level index is at the lowest level. Combining all three Series and Element Descriptors leads to a medium Overall Series and Element Risk.



Supplemental Information Table. Definitions of nodes and states used in the Bayesian network model of archival 35mm B&W film priority ratings (see text for explanation) for National Archives and Records Administration (NARA), USA. The case file database mentioned is referenced in the main text.

Submodel, Node, States	**Definition**
Submodel: Series & Elements	
Node: Series	Series is how NARA organizes records
States: 111-ADC 330-DVIC 306-. 428-NPC 111-LC 306-LN 018-CS 342-USAF 127-G 319-IDF 242-MID	These are the 11 series that are most frequently requested based on the Labwork database. They represent about 70% of the total requests over the past decade or so, as of 2023.
Node: Element	Element or format of the film; these are black and white (B&W), 35mm, with no soundtracks.
States: DNC DNS DRC DRS FGMC FGMS MPS MPPC MPPS ONC ONS ORC ORS PPC PPS	Source of element type listing: NARA (2021, Appendix 2) - https://www.archives.gov/files/preservation/formats/pdf/nara-motion-picture-technical-processing-guidance.pdf - Page 11

Submodel: Series and Element Descriptors	
Node: Series and Element % Unique P Copies	Percentage of unique preservation (P) copies in the HMS catalog.
States: 0 >0 to <30 >=30 to <50 >=50 to <70 >=70 to 100	State value ranges were defined via visual inspection of the frequency histogram of the distribution among Series and Elements, and what seems reasonable as judged by NARA archivists and lab staff.
Node: Use Based on Labwork Percentage	Estimate of "use" based on the percent of total preservation entries in Labwork across all Series and Elements. Labwork data covered periods from late 2005 to early 2023.
States: 0 >0 to <0.0002 >=0.0002 to <0.5 >=0.5 to <1 >=1 to <2.5 >=2.5	State cutoff values are determined based on discretization from NARA's Holdings Management System (HMS) database, modified by reasonable values usable by NARA archivists and lab staff. The zero level is a hard zero; there are no Series/Element combinations of that type. The <0.0002 level represents Series and Elements which do exist but which have not come into the lab over the past ~17 years. For the two highest level states, the percent of use by series generally follows a Pareto distribution.
Node: Acidity Level Index	Based on estimate of acidity distribution curve shape of the number of selected records at various acidity levels, as informed from Labwork.
States: Below average (1 to <1.5) Slightly below avg (1.5 to <2) Average (2 to <2.5) Slightly above avg (2.5 to <3) Above average (3 to < 3.5) High (3.5 to <4) Very high (>=4)	The average Acidity Level Index score was defined as 2, representing the average over all combinations of Series and Elements. Other levels of the Acidity Level Index were based on deviations from this average score, based on frequency histogram distributions. The actual acidity measurements on which this scale is based were made with AD strips from the Image Permanence Institute. AD strips give results in a range of 0 to 3 in steps of 0.5, but these ratings should not be confused with the Acidity Level Index scores used in this node.
Submodel: Archival Values	
Node: Archival Significance	"Significance" is defined based on four factors pertaining to government and historical research significance: - intrinsic significance of the physical item (film) - government significance of content - historical research significance of content - rights and interests significance These values have been adapted for motion picture film based on NARA (2023).

States: 1 2 3	1-3, higher values denote greater significance, as follows: 1: Low significance. Records were the product of minor or routine government functions and have limited informational or evidential value. These are often series created by functions in the course of administering a business process or a routine program such as providing individual entitlements, carrying out registrations (such as draft cards and passports), issuing licenses, or providing services. Other series held by NARA or published sources may contain the same amount of information. The records were usually created after 1950. 2: Moderate significance. Records may have low or limited value as artifacts, but contain important information about, or were produced as a result of, major government processes and functions. They provide evidence of the basis for substantive functions performed by a federal agency, commission, or court, and possess evidentiary value for documenting policy formation and government decisions at leadership levels. The information contained in the records is generally considered to be unique to the records, with no other known sources having the same range and type of information or reflecting the way it is organized. Age may also be a factor and records that otherwise may be considered of low significance could be designated as having moderate significance primarily because they date to a period when few records survived or before the National Archives and Records Administration was created. Generally, all records created before 1930 should be considered as moderately or highly significant regardless of other characteristics. 3: High significance. Records that have exceptional significance due to critical informational and/or evidentiary value. Significant harm to the Federal Government would occur if lost. The records may be considered international or national historical icons or treasures, unique or irreplaceable, and have few peers, or they may be viewed as works of art in addition to sources of information. These records may also provide exceptional documentation of functions and activities at the highest levels of government. Appearances of historic figures and great age (generally NARA holdings created before 1930) may be an important determinant of high significance independent of most other characteristics. Highly significant records often have a high monetary value and could be attractive targets for theft. Significant harm to the Federal Government would occur if these records are lost. Source: NARA (2023)
Node: Archival Uniqueness	"Uniqueness" is defined based on 4 criteria: - This is the only source of information. - The same information is not available in other NARA holdings. - The information is not available in other publicly available sources (published or not). - If other sources exist, this source provides additional information not found in any other source.

States: 1 2 3	1-3, higher values denote greater significance, as follows: 1: Low uniqueness. Recordings of commercial broadcasts, press files. 2: Moderate uniqueness. Permanent records of materials acquired by an agency from a third party, or joint private/public partnerships; also, titles with record copies on film in multiple series. 3: High uniqueness. Records that are unique and irreplaceable and have exceptional historical significance due to documentation of a major event. Source: NARA (2023)
Node: Archival Value	Archival Value is a numerical integer sum of the rating levels of Archival Significance and Archival Uniqueness.
States: 2 3 4 5 6	As above.
Submodel: Priority and Risk Ratings	
Node: Series and Element Acidity Risk	A deterministic combination of Acidity Element and Use Based on Labwork Percentage.
States: Low acidity risk Medium acidity risk High acidity risk Highest acidity risk	Series and Element Acidity Risk levels are higher with greater Acidity Level Index and higher Use Based on Labwork Percentage.
Node: Overall Series and Element Risk	A deterministic combination of Series and Element Acidity Risk, and Series and Element % Unique P (Preservation) Copies.
States: Low risk Medium risk High risk Highest risk	Overall Series and Element Risk levels are higher with higher values of Series and Element & Unique P (Preservation) Copies and greater Series and Element Acidity Risk.
Node: Overall Series and Element Priority Rating	Overall Priority Rating is the joint risk analysis arising from Archival Value and Overall Series and Element Risk levels.
States: No action needed Low Medium High Highest	Overall Potential Priority Rating levels are higher with greater Archival Value and greater Overall Series and Element Risk.

Submodel: Ancillary Information	
Node: Series Percentage Unique P Copies	Percentage of unique Preservation (P) copies of all Element categories for a specific Series. This is provided in the model for record-keeping but does not directly influence the Overall Priority Rating.
States: 0 to <1.9 >=1.9 to <6 >=6 to <17.5 >=17.5	State value ranges were defined via the discretization algorithm in Netica for equal numbers of cases in each state, using the case file database.
Node: Series Count of Unique P Copies	Sum of unique Preservation (P) copies of all Element types for a specific Series. This is provided in the model for record-keeping but does not directly influence the Overall Priority Rating.
States: 0 to <470 >=470 to <1000 >=1000 to <2900 >=2900 to <4420 >=4420	State value ranges were defined via the discretization algorithm in Netica for equal numbers of cases in each state, using the case file database.
Node: Series and Element Labwork Count	This is the number of Preservation entries in the Labwork. This is provided in the model for record-keeping but does not directly influence the Overall Priority Rating.
States: 0 >0 to <9 >=9 to <29 >=29 to <90 >=90	State value ranges were defined via the discretization algorithm in Netica for equal numbers of cases in each state, using the case file database.
Node: Series and Element Dimension Not Available	Number of entries of Series and Elements where the dimension (film gauge) is not listed in HMS (Holdings Management System of NARA). It is possible that some entries are 16mm and not 35mm. This is provided in the model for record-keeping but does not directly influence the Overall Priority Rating.
States: 0 >0 to <25 >=25 to <75 >=75 to <250 >=250	State value ranges were defined based on utility for the NARA archivists and lab staff. Values are assigned from a practical viewpoint, that is, entailing what would imply a manageable effort to track down the missing data.
Node: Series and Element Count of Unique P Copies	Number of unique P (Preservation) copies in HMS catalog among all elements; this pertains to P that does not have an Intermediate (I) copy. Note: since dimension is unknown for some entries, the total may include some 16mm copies and not the targeted 35mm. This is provided in the model for record-keeping but does not directly influence the Overall Priority Rating.

States: 0 >0 to <10 >=10 to <100 >=100 to <500 >=500 to <1500 >1500	State value ranges were defined based on what seems appropriate for NARA archivists and lab staff. The intervals denoted are easily manageable.
Submodel: Case Number	
Node: Case Number	Number of the case (row) in the case file database.
States: run 1 to 165	As above. For tracking purposes.

References

NARA. 2021. National Archives and Records Administration (NARA) Motion Picture Technical Processing Guidance. https://www.archives.gov/files/preservation/formats/pdf/nara-motion-picture-technical-processing-guidance.pdfhttps://www.archives.gov/files/preservation/formats/pdf/nara-motion-picture-technical-processing-guidance.pdf

NARA. 2023. NARA 1571 Supplement 3, Determining the Significance of NARA Holdings, January 17, 2023. https://www.archives.gov/foia/directives/nara1571https://www.archives.gov/foia/directives/nara1571