A Bibliometric Approach to Characterizing Technology Readiness Levels Using Machine Learning

Jehan Dastoor*, Heying Zhang†, Michael Balchanos‡, and Dimitri Mavris§
Georgia Institute of Technology, Atlanta, Georgia, 30332

As cislunar space becomes more accessible to national space agencies and commercial entities, there is a constant need to improve the way in which space missions are planned, and development progress is tracked. A technologies stage of development, which is related to mission budget and schedule, is typically quantified using technology readiness levels (TRL). The process of determining TRL is often long and laborious, and requires the use of subject matter experts. As a part of the Georgia Institute of Technology Cislunar Architecture Initiative, this work serves to develop the early stages of an environmental scanning approach to maturity assessment that allows for the automatic determination of a technologies TRL using machine learning ordinal regression techniques with bibliometric factors. The bibliometric factors considered were: scientific publications, National Science Foundation awards, patents, and NASA Spinoff articles. Annual data on these factors was collected for 31 technologies between 1995-2015 using public APIs, and S-curves fit to the data to estimate each technologies point in the development cycle. The final model performed with an R^2 of 0.817, 0.812, and 0.567 on the training, validation, and test data, respectively. Additionally, a better performing model to classify a technologies technology life cycle phase was created and drawbacks to this approach discussed.

Nomenclature

AHP	=	Analytical Hierarchy Process
API	=	Application Programming Interface
a	=	Logistic growth diffusion point
BIMATEM	=	Bibliometric Methods for Assessing Technology Maturity
b	=	Logistic growth velocity of diffusion
k	=	Logistic growth upper limit
MFE	=	Model Fit Error
MRE	=	Model Representation Error
NSF	=	National Science Foundation
R^2	=	“R-squared” or the coefficient of determination
$r(t)$	=	Logistic or S-curve function
TLC	=	Technology Life Cycle
TRA	=	Technology Readiness Assessment
TRL	=	Technology Readiness Level
TRL_{adj}	=	TRL adjusted on a cardinal scale
t	=	Time
t_0	=	Starting time
USPTO	=	United States Patent and Trademark Office

*Graduate Research Associate, School of Aerospace Engineering
†Graduate Research Assistant, School of Aerospace Engineering
‡Research Engineer II, School of Aerospace Engineering, AIAA Member
§Regents Professor, School of Aerospace Engineering, AIAA Fellow

I. Introduction

In GAO's 2021 release of its Assessment of Major Projects, NASA was pointed out to have a 31.1% cumulative cost overrun of $9.6 billion, and a cumulative schedule overrun of 19.7 years (as can be seen in Figure 1) [1]. The deficiency to control and predict progress in major projects has lasting and detrimental effects on NASA's acquisition portfolio which ultimately limits the amount, type, and control over funding and projects [2]. As interest in cislunar space grows, defense, science, and defense stakeholders are anticipated to attempt more cooperative and complicated projects. In order to meet this growing complexity, it is imperative to develop new and improved methods to better forecast mission budget and schedule and track mission development.

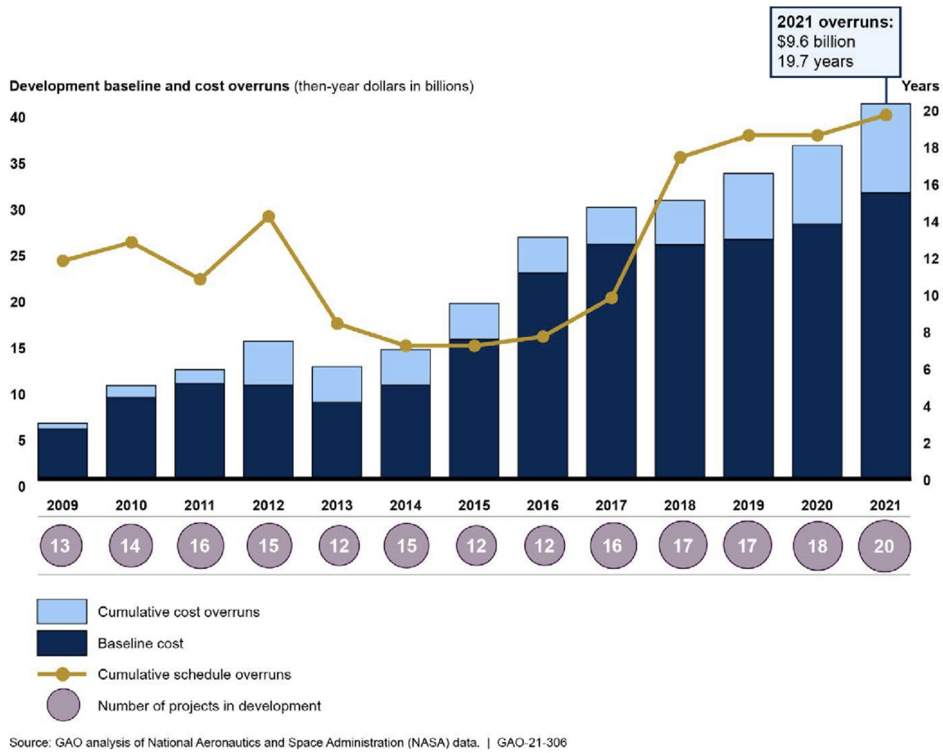


Fig. 1 Cost and schedule overruns for major NASA projects [1].

Traditionally, these forecasts have been carried out by looking at the maturity of required technologies, which provides a glimpse into how much future effort is required to develop a technology to the point where it can be used for a mission [3]. The maturity of a technology is typically described using the technology readiness level (TRL), where technologies with low TRLs require significantly more investment and development time than technologies with high TRLs that are said to be mature. As such, it is often a major focus of mission planners to determine TRL quickly and accurately.

The typical technology readiness assessment process used to determine TRL is often time consuming and involves querying a group of subject matter experts until a consensus is reached [3]. This presents a number of drawbacks. For example, the availability of subject matter experts is often limited [4]. This is made especially difficult for complex missions and large projects where a large number of technologies must be assessed, and collecting a group of subject matter experts on every technology is neither realistic or feasible within the time constraint. Personal and organizational biases, and expectation pressure may also influence the final assessment [4]. An additional drawback is that the technology readiness assessment process takes a significant amount of time and only provides a snapshot of a constantly evolving technology. Thus, multiple assessments are often required throughout a project's life cycle to track development goals and provide better forecasting estimates [4]. These requirements make it difficult, expensive, and time consuming to rely on subject matter experts for this process.

As a result of these problems, a significant amount of research has been done to try to automate the process of determining TRL. One such method, known as extrapolation, involves using previously collected maturity data to extrapolate to future points by fitting the data to a chosen growth curve [3]. Young, DeBecker and Modis, and other authors have looked into how to select the correct growth function, and how to compute uncertainties in these growth parameters [5, 6]. Another more recent approach, proposed by Watts and Porter, has been to use bibliometric factors to assess TRL [7]. Bibliometric assessments for technological maturity involve the analysis of books, articles, publications, and other research trends to extract where in the development cycle a technology currently exists. For example, technologies with a high maturity likely have patents associated with them, whereas lower maturity technologies have not reached a developmental point where patents can be granted [7]. Authors such as René Lezama-Nicolás et al., and Safa Faidi have used scientific publications, patents, and news articles to determine TRL using an acceptance threshold approach [8, 9]. Cauthen et al. recently proposed a method to characterize technologies in the emerging (TRL 1-5) or growing (TRL 6-7) stages of development using an artificial neural network classifier trained on scientific publication data [10]. While this recent advancement is exciting, it only makes use of one bibliometric factor to distinguish between stages of development, however risk, budget, and schedule assessments are made at a single TRL level.

As a part of the Georgia Institute of Technology Cislunar Architecture Initiative, this work aims to provide an initial attempt at creating an automated framework for determining TRL through the use of machine learning techniques with bibliometric factors. This will enable mission planners and forecasters to provide maturity, budget, and schedule estimates quickly, without the need for subject matter experts, allowing parametric forecasting simulations to be run easily, and real-time assessment of project progress. Additionally, the framework provided is easily scalable to assess a large number of technologies in the case of a large project. Ordinal regression techniques used to create the model have been discussed in the following sections, and model and classification performance metrics are provided.

This paper is organized as follows. In Section II we provide an introduction and literature review of TRL, the technology life cycle, and methods for determining TRL. In Section III we introduce the proposed framework, define the bibliometric factors used, provide example S-curve fits to these factors, discuss the chosen data collection techniques, and introduce the topic of ordinal regression and the naive regression model used in this study. In Section IV, we provide results for the trained regression model including the coefficient of determination, R^2, actual versus predicted plots, model fit and representation errors, and confusion matrices for the classification of a technologies TRL level. We conclude by providing observations that can be used to improve the models predictive capabilities for future spirals of this research.

II. Background

A. Technology Readiness Level

The technology readiness level (TRL) scale is an ordinal scale used to assess the maturity of a technology, ranging from 1-9, with 1 indicating no maturity and 9 indicating a fully mature and tested technology [11]. The scale has seen uses in a variety of different applications from space mission planning to software evaluation. Since TRLs introduction in 1989, TRL has seen widespread acceptance as the standard metric used in technology maturity assessments for a variety of reasons [12]. TRL being a standardized metric allows for the comparison of the maturity of different and unrelated technologies. Additionally, TRL has been shown to predict budget, schedule, developmental risk, and a number of other factors associated with the process of maturing a technology [13]. This makes TRL an incredibly useful metric to quantify during a technologies initial development so that budgets can be correctly allocated and timelines properly disclosed. Notably, the assessment of a technologies TRL level must be done continuously throughout its development in order to ensure the most accurate predictions. The overall maturity process of a technology is known as its life cycle.

B. Technology Life Cycle and the S-Curve

The life cycle of a technology describes the stages in the life of a technology, from the beginning conception to the final replacement with a newer more advanced technology. The technology life cycle (TLC) indicator model was first proposed by Watts and Porter and is typically described using 4 stages: the emerging stage, the growing stage, the mature or saturation stage, and lastly the declining phase [7, 9]. Each stage is associated with a range of TRLs, with the emerging stage corresponding to TRL 1-5, the growing stage to TRL 6-7, and the mature stage to TRL 8-9. The declining phase is not shown on the current TRL scale. This life cycle has been shown by various authors to follow the shape of an S-curve over time (or logistic growth curve) relative to development cost, number of patents, number of published scientific or engineering articles, and a number of other factors [7, 14–16]. The S-curve is described by the following equation:

$$r(t) = \frac{k}{1 + ae^{-b(t-t_0)}}, \tag{1}$$

where $r(t)$ is the S-curve function (or the logistic growth function), k is the upper limit of growth, a is the initial stage of diffusion, b is the velocity of diffusion, t is time, and t_0 describes where in time the curve begins [9]. Note that the S-curve is commonly plotted against time (and will be in the rest of this paper), however it has also been plotted against TRL by Terrile et al., with the steepest increase seen around TRL 6-7 [15]. The shape of the curve provides valuable information, that can be used to predict cost and schedule at the next stage of development, not just the current state of the technology. As such, understanding where a given technology is located on the S-curve is of great importance for decision makers seeking to develop investment strategies, predict future budgets, or solely assess TRL.

C. Technology Readiness Assessment

The process of determining TRL is known as technology readiness assessment (TRA) and can be carried out in a variety of ways [4]. The most common way for assessing TRL is known as Delphi, which aims to find consensus amongst a group of experts by asking them yes or no questions that are answered with the use of supporting evidence [3]. The process aims to be as objective as possible, however the choice of questions, the evidence available, and the person answering the question all have great influence over the final outcome. As such, results from a Delphi are often thought of as biased and subjective [17]. Other TRA methods such as extrapolation exist. This method uses historic data on a technology to extrapolate how mature it may be in the future, typically using S-curves or the Gompertz function [3]. However, this assumes that previous data on the technology is readily available, and that a portion of a technologies history contains all the necessary information to forecast its development, which is not always the case.

A more recent method is that of environmental scanning which utilizes the fact that technological change typically happens following a common sequence of steps [3]. A typical sequence might include: theoretical proposal, scientific findings, engineering reports, developing prototypes and procuring patents, and lastly commercial introduction [3]. By observing one phase in this sequence, one might be able to predict when the other phases will occur. To do this, a thorough search through scientific, trade, and business literature must be conducted to find mentions of the technology to assess its current stage in the sequence. This is typically done using bibliometrics, which is the statistical analyses of text based material. Bibliometric methods have been used by authors, such as René Lezama-Nicolás et al. and Safa Faidi, to estimate TRL and are typically called Bibliometric Methods for Assessing Technology Maturity (BIMATEM) [8, 9]. Phases of this typical sequence can then be matched to TLC stages to begin characterizing a technologies TRL. The model proposed by Watts and Porter provides example bibliometric indicators and connects these to parts of the S-curve and their corresponding TLC phase as shown in Figure 2.

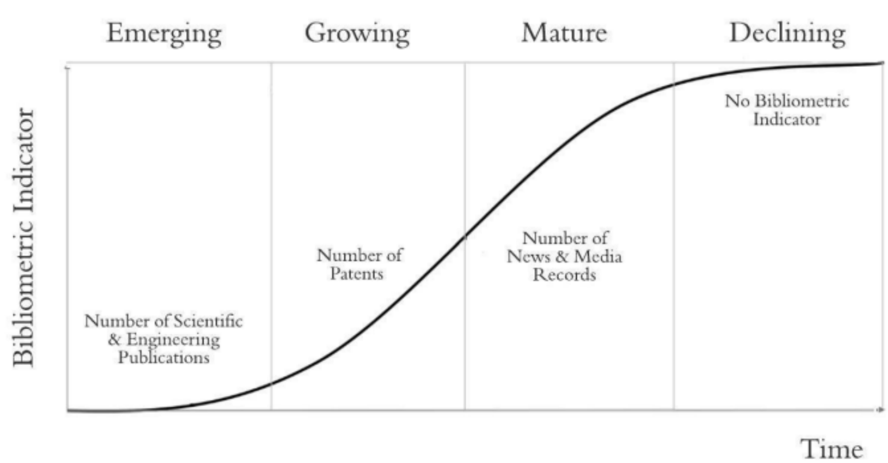


Fig. 2 Bibliometric indicators associated with each TLC stage [9].

There are a number of advantages associated with the environmental scanning approach, such as the removal of the need for subject matter expert input, and the removal of subjectivity by using a more rigorous statistical based approach that capitalizes on existing literature and databases. Additionally, bibliometrics and machine learning also allow for the possibility of automating the TRA process using software based data collection and evaluation techniques, resulting in a significant improvement in the time required to carry out TRAs. However, the main disadvantage is that bibliometric analysis relies heavily on the use of provided keywords to search databases and acquire relevant information. This process of keyword selection is often difficult, and can require prior knowledge on the technology being considered for the best results. Additionally, the same selection of keywords can apply to a number of different technologies making it difficult to ensure papers on the intended technology are being gathered. Overall, the possibility of automating the TRA process makes bibliometrics an interesting field to explore.

Current BIMATEM techniques use the S-curve model during the emerging and growth stages for scientific papers and patents. However, when looking at news and media records these techniques tend to use the hype-type evolution model, which is an extension of the S-curve [9]. The overall process can be seen below in Figure 3.

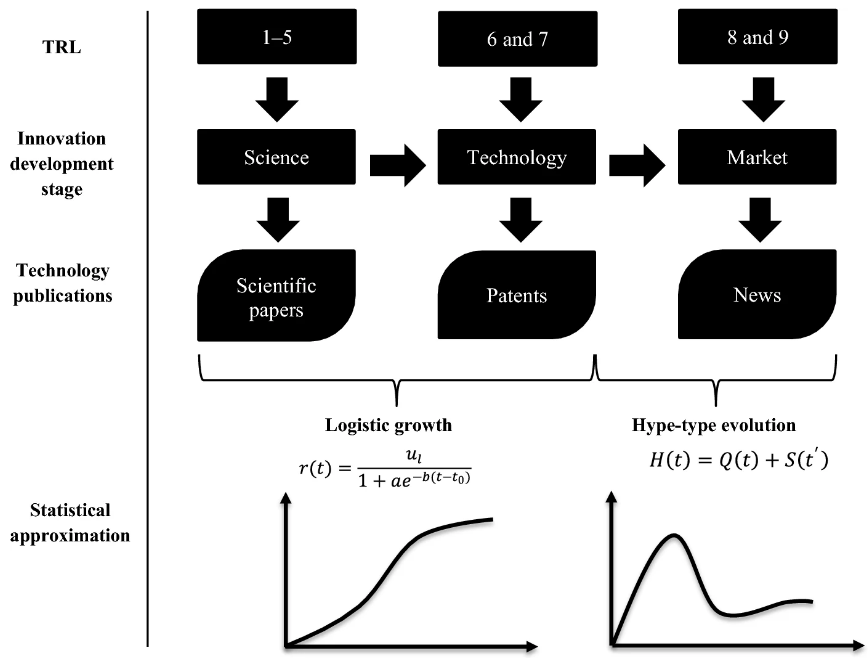


Fig. 3 Bibliometric factors corresponding to each TLC stage and their respective statistical approximations [8].

Note that while a technology has an overarching maturity S-curve associated with it, bibliometric factors for that technology may also have their own individual S-curves. These individual curves are more valuable when determining TRL as they provide information on the maturity of a technology relevant to each metric, whereas the overall maturity of a technology is difficult to quantify without already knowing the TRL. The position of the technology on each of these curves can then be passed to a model or decision making technique to determine the TRL.

There have been a variety of models proposed that utilize this bibliometric information to predict TRL, such as: an Analytical Hierarchy Process (AHP) [18], or an acceptance threshold model where surpassing the threshold for each metric indicates a specific TLC stage [9]. Another approach, which is utilized in this paper, is to create a machine learning model which determines which maturity group (corresponding to the TLC stages) a technology is in. This approach has been used successfully by Cauthen et al. in the past to classify technologies as being either in the emergent or growth stage of development using only scientific publications [10]. However, more metrics can be considered to get a more holistic picture of a technologies maturity, whereby at least one metric should be used to represent each TLC stage. Utilizing a variety of metrics, it is possible to predict any TLC stage. The distinction between different maturity phases is not always clear, and different technologies develop differently depending on their popularity. Accounting for this non-linearity will require the use of machine learning.

III. Methodology

The method outlined here provides a way to automate the TRA process by utilizing bibliometric techniques, without the necessity for subject matter experts. The proposed TRA process can be seen in Figure 4 below. Note that any method can be used to collect data on the relevant bibliometric techniques (such as manual searching), however the API method shown in the Figure was the method used to automatically collect data for this study. Other methods will likely forgo some aspects of automation, yet might provide greater accuracy. This decision is left up to future users of this method.

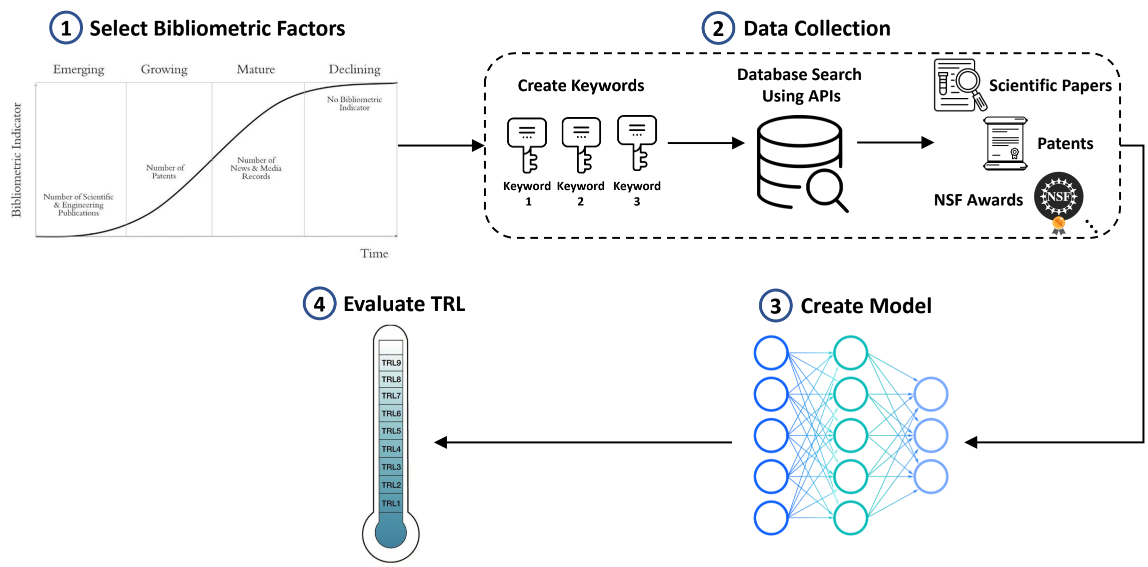


Fig. 4 Framework for predicting TRL from selected bibliometric factors.

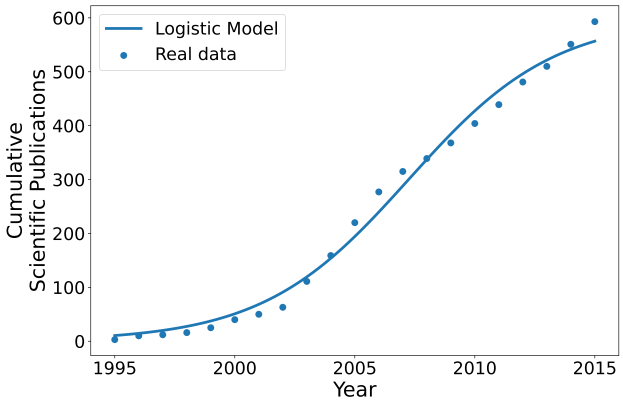
The process of data collection was automated using publicly available APIs to search through relevant databases for each bibliometric factor. The specific factors chosen and their relevant APIs will be discussed in the following part of this section. Note that for each metric that required fitting with an S-curve the Python library SciPy was used [19].

A. Bibliometric Factors

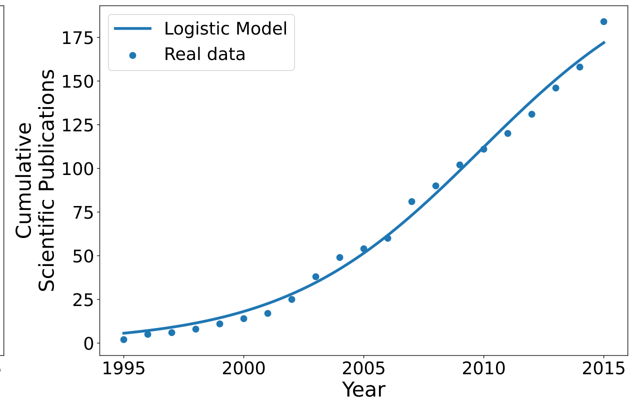
The first part of the method involves selecting bibliometric factors that will be used as inputs to the TRL prediction model. A variety of factors can be used depending on the stage in the TLC a technology is at. The factors used in this paper have been described below, although it should be noted that a variety of other factors can be used if desired. Note that the declining phase is not discussed further as this phase is not represented on the standard TRL scale of 1-9.

1. Emergent Phase

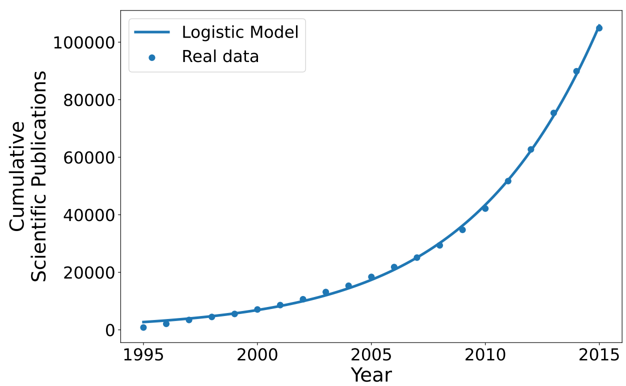
The emergent phase corresponds to TRL values between 1-5, and is indicative of a technology that is still in the conceptual phases of its development and hence, is being largely developed by the scientific community [7]. As such, the current state of scientific and engineering literature on the technology provides a good indication as to whether the technology is in the emerging phase. The number of papers published over time is initially low as few scientists are looking into the basic principles of the technology, and then begins to rise as feasibility is proven and more scientist begin to create proofs of concept. Eventually, development of the technology is passed onto commercial firms who are not likely to publicize information on the technology in order to maintain their proprietary data and knowledge. This trend resembles that of an S-curve. An example of this can be seen for solar sail propulsion (TRL 5), deep-space position systems (TRL 9), and photovoltaic solar cells (TRL 9) in Figure 5 below [20].



(a) Solar sail propulsion.



(b) Deep-space position systems.

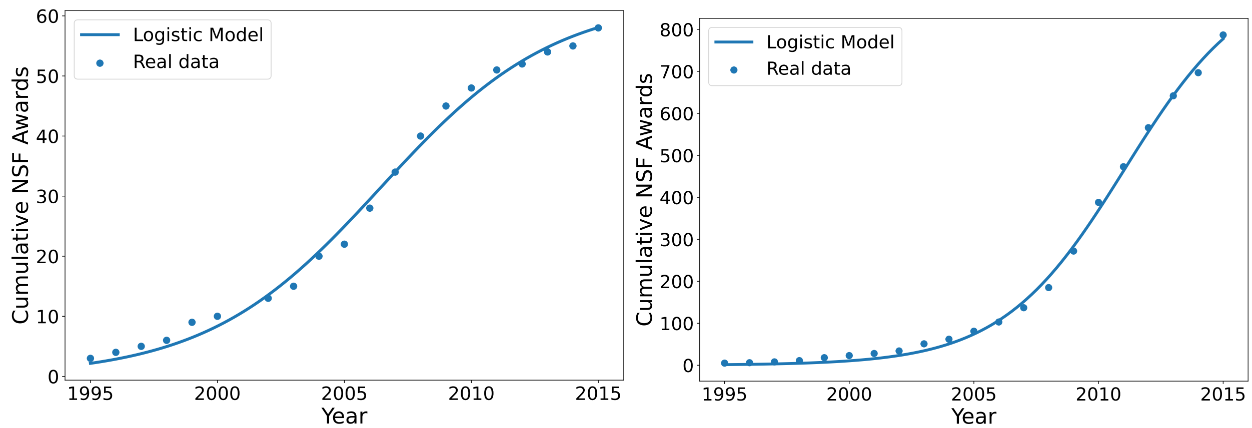


(c) Photovoltaic solar cell.

Fig. 5 Example of the cumulative scientific publication counts for three technologies between 1995 and 2015 (data acquired from the Scopus database [21]).

Looking at Figure 5, it cannot easily be determined which technology is more mature. For example, even though deep-space position systems and solar cells have a TRL greater than solar sail propulsion, it appears as if the technology is earlier on its S-curve as many scientists are still researching how to make positioning systems and solar cells more efficient. Therefore, only looking at one metric is not a good indication as to the state of maturity of a technology.

Another good indicator for technologies in the emerging stage (as suggested by Faidi [9]) is the number of National Science Foundation (NSF) awards granted over time for the development of the given technology. Awards offered by entities, such as the NSF, are typically provided to support fundamental research for technologies that are in the early stages of development and would not be funded otherwise due to the high probability of failure during this stage [22]. The number of awards granted also follows an S-curve as awards are typically provided for research that is later published in journals, and hence this information is likely correlated with the previous metric. An example of this can be seen for deep-space position systems and photovoltaic solar cells in Figure 6 below (note that not enough awards were found for solar sails to create an accurate S-curve and so this plot has not been shown).



(a) Deep-space position systems.

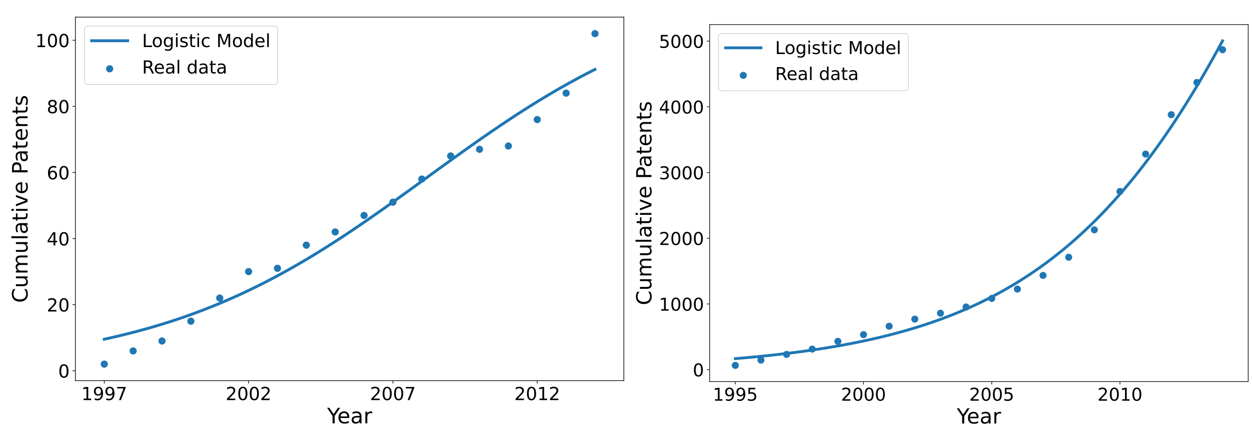
(b) Photovoltaic solar cell.

Fig. 6 Example of the cumulative National Science Foundation (NSF) award counts of two technologies between 1995 and 2015 (data acquired from the NSF database [23]).

The S-curves in Figure 5 and Figure 6 are relatively similar, highlighting the correlation between these two metrics. Hence, these metrics can act as a backup for each other. For example, solar sails do not have enough awards to produce an accurate awards S-curve, but there are enough scientific publications to produce an S-curve based on publications.

2. Growth Phase

The growth phase corresponds to TRL values between 6-7, and is typically characterized by a sharp rise in the maturity of a technology as initial proofs of concept have been validated, and inventors and industry have begun to start developing systems for real applications [7]. This is the point where proprietary data starts becoming more prevalent and stakeholders are likely to begin patenting their inventions in order to ensure larger profits in the future. Thus, a valid bibliometric factor to track is the number of patents for a given technology over time. Initially, the number of patents rises as a new innovative concept encourages people to develop new products. Eventually as more patents are put forward, the number of future possible patents for the given technology decreases until there are no more possible patents or there is a better alternative technology available. Hence, similar to scientific papers, the trend in number of patents over time will likely resemble an S-curve. An example of this can be seen for deep-space position systems and photovoltaic solar cells in Figure 7 below.



(a) Deep-space position systems.

(b) Photovoltaic solar cell.

Fig. 7 Example of the cumulative number of patents for two technologies between 1995 and 2015 (data acquired from the United States Patent and Trademark Office using the PatentsView API [24]).

Again, from this Figure solar cells seem as though they are still in the early stages of the growth phase. It should be noted that this is likely due to the fact that many different kinds of solar cells exist. Overall, it may seem as though solar cells are still in the early stage of development, but it is possible that some kinds of solar cells have S-curves representative of a mature technology whereas others have S-curves representative of an early developing technology. As such, care must be taken to only collect data on the exact version of the intended technology.

3. Mature Phase

The mature phase (or saturation phase) corresponds to TRL values between 8-9, and is described as the final stage in the maturity of a technology where the technology has been fully developed and must now be validated in its intended environment. At this point, the technology (if popular) is likely being discussed by news outlets for its utilization (or planned utilization) in real systems. If this is true, one can determine the number of news articles published over time to estimate maturity. This time series would follow the hype-type evolution shown in Figure 3. However, if the technology is not used on public facing systems or is not popular, the number of news articles published on the technology will be reflective of this and it will likely be difficult to produce an accurate S-curve. Additionally, news APIs are typically not free and do not date back far enough to create a sufficient time trace. Therefore, another method is needed to either supplement news data or provide an alternative if the technology is not picked up by the news.

Notably, at this phase of development the technology is being developed by commercial entities and is being manufactured at some scale. As such, one could look to financial information such as public venture capital investments or generally, commercial uses for this technology. Looking at the number of commercial uses for this technology currently, or the time series of commercial uses can provide valuable information on the adoption of this technology, where the greater the adoption, the more likely the technology is to be mature.

Therefore, we have a selection of four bibliometric factors for this study: scientific publications, NSF awards, patents, and commercial uses or availability of a technology.

B. Data Collection

Data for each of these factors can be collected utilizing publicly available APIs. These APIs allow for the automated searching of a public database and provide the user the ability to collect all entries that match a given search query. The data collected is very sensitive to the query used and hence, care must be taken in designing these queries. Designing these queries, depending on the complexity of the technology, may require the input of subject matter experts. However, the search is ultimately limited by the capability of the chosen API and the search fields it provides. This makes the selection of each API, and relevant database, an important task that must be thoughtfully considered.

1. Scientific Publications

To collect information on scientific publications, the Scopus API connected to the Scopus database by Elsevier was chosen. The Scopus database contains over 78 million items from over 5,000 different publishers dating back to 1966 [21]. As such, it is an excellent database to model the number of papers over time as it provides a large collection of papers, from a variety of sources, over a wide range of dates. Additionally, the Scopus API allows the user to query for papers published between specific dates, and for specific keywords in the title and abstract of the paper. The query can also contain logic statements such as "OR" and "AND" to chain together keywords and create a more specific search for a given technology. This allows for advanced searches to be written in plain text, making it easier to write complicated search queries.

The Scopus API was used to collect the number of scientific publications per year that matched certain keywords. An S-curve was then fit to this data and the normalized value of the S-curve (i.e. $r(t)/k$) at the current point in time was provided as an input to the TRL prediction model. A normalized value of one indicates full maturity, whereas a value of zero indicates no maturity for that metric. For example, solar sail propulsion was found to have a normalized value of the S-curve in 2015 of 0.947 or 94.7% mature according to the number of publications written (this can be seen to match the results shown in Figure 5a where solar sail propulsion is near the end of the S-curve). The code to collect information from the Scopus database was adapted from Safa Faidi's thesis [9].

2. National Science Foundation Awards

Information on awards given for fundamental research was collected using the National Science Foundation (NSF) awards API from the NSF awards database [23]. This database contains information on awards granted since 1989, providing a large enough time window to collect data to fit an S-curve. Additionally, NSF accounts for approximately 20% of all federal support provided by the US government to academic institutions for basic research [23]. Therefore, the database is comprehensive enough to approximate the number of monetary awards being granted for fundamental research for a given technology. It should be noted that this API does not provide as much flexibility as the Scopus database in how to define a search based on keywords. The user is still able to search for a phrase exactly and filter results by the date the award was granted, but the only logical statement available is the "OR" statement indicated by a "+" sign. This makes it more difficult to create complex search queries.

Similar to scientific publications, the NSF awards API was used to collect information on the number of awards granted per year for a given technology. An S-curve was then fit to this data and the normalized final value of the S-curve provided as an input to the model. For example, the normalized value of the S-curve in 2015 for deep-space position systems was determined to be 0.942 or 94.2% mature according to the number of NSF awards provided, which is reflected by Figure 6a showing a nearly fully formed S-curve.

3. Patents

Patent information was collected using the PatentsView API which connects to the United States Patent and Trademark Office (USPTO) database to collect information on US patents. The database contains over 5 million US patents from over 1.2 million inventors, dating back to 1976 [24]. Similar to the Scopus API, the PatentsView API provides users the flexibility to use "OR" and "AND" statements, and to search by date. However, these queries are constructed in a more complicated way than the Scopus API and are no longer written in plain text.

The PatentsView API was used to track the number of patents related to a given technology over time in order to generate a best-fit S-curve. The normalized final value of this S-curve was then provided as an input to the model. For example, photovoltaic solar cells, shown in Figure 7b, had a patent maturity of 88.2%. The code to collect patent information using this API has been adapted from Safa Faidi's thesis [9].

4. Commercial Uses

Lastly, commercial use information is difficult to come by without access to paid for financial and venture capitalist databases. As an alternative to this, the NASA Spinoff database and API were used. The Spinoff database highlights over 2,000 NASA technologies since 1976 that have benefited life on Earth in the form of commercial products and services [25]. Note that it is assumed here that technologies used in commercial products must be of a high maturity to be economically viable and hence, these technologies are likely in the mature phase of the TLC. On the downside, the Spinoff database does not allow the user to search by year, only by one specific keyword or phrase. Therefore, better databases and APIs to predict mature technologies will likely need to be used in the future.

Unfortunately, since this API does not allow for querying by date, one can only acquire the total number of Spinoff products and services related to a given technology. Since technologies vary in popularity, this value was divided by a technologies combined total number of scientific publications and patents to get a more reasonable value that can be compared across technologies. This new value will be called the Spinoff norm in the rest of this paper. The combined total is a good indication of popularity as the excitement surrounding a technology influences the number of scientists interested in conducting research and the number of inventors seeking to patent the technology and vice versa.

C. TRL Prediction Model

The TRL prediction model takes these bibliometric factors as an input and returns the predicted TRL or TLC. Since TRL is an ordinal scale, care must be taken in choosing a machine learning model. One approach is to use ordinal regression techniques, which are supervised machine learning techniques, where the objective is mixed between regression and classification [26]. These techniques can be broken down into three major categories: naive approaches, ordinal binary decomposition's, and threshold models, as can be seen in Figure 8 below.

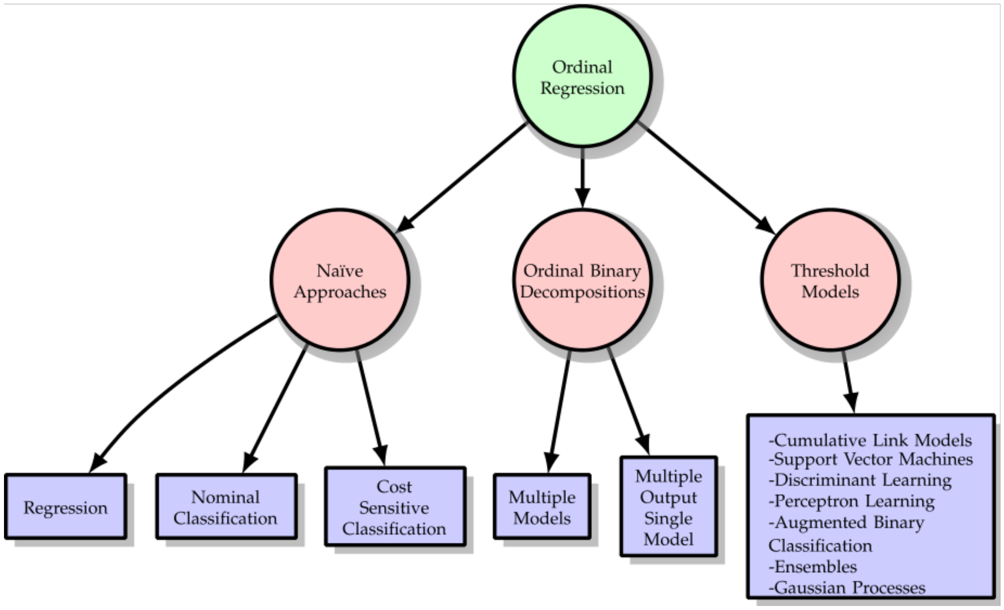


Fig. 8 Different possible machine learning approaches to ordinal regression [26].

For this paper, a naive regression approach was used to provide a proof of concept for using multiple metrics to predict TRL. However, other authors such as Cauthen et al. have used binary classifier algorithms to the same effect [10]. In order to use a regression approach, the TRL scale was converted to a cardinal scale using Conrow's TRL adjusted, $\mathrm{TRL_{adj}}$, curve fit which was developed using an AHP [27]:

$$\mathrm{TRL_{adj}} = 0.346 + 0.012\,\mathrm{TRL}^3. \tag{2}$$

These adjusted TRL values were assumed to be continuous and then regressed against using a neural net in JMP. The actual TRL value can then be determined by taking the inverse of Eq.2 and rounding to the nearest integer. Some accuracy will likely be lost in the rounding process, which is why binary decomposition and classification models are likely to perform better. However, these models require a greater amount of training data in order to provide accurate predictions. In future spirals, the modeling approach provided here will likely be improved as more data is collected.

IV. Results

The NASA 2015 Technology Roadmap [20] was used to select 31 different technologies and acquire their respective TRLs to be used in this study with a 19-6-6 split between training, validation, and test data, respectively. Since the 2015 roadmap was used, each database was searched from 1995-2015, with the S-curve being evaluated at the last year 2015. The list of technologies and the search queries used for each can be found in the Appendix in Table 1. Note that these queries were not produced with the help of subject matter experts, and therefore may not accurately describe the technology. Ideally, more technologies would have been used to create a better model, however due to time constraints only 31 technologies were used.

These search queries were used by the APIs described in the previous section to collect relevant bibliometric factors: S-curve values for scientific publications, NSF awards, and patents, and also the number of Spinoff articles, and the Spinoff norm. A summary of the collected data can be seen in the Appendix in Table 2. If no results were found, the S-curve value was set to zero for that metric as it is assumed that the technology must not be in that specific stage of development (indicated by that metric) in order for there to be no results. Using these results, a regression based neural network was trained in JMP to predict the adjusted TRL value. The best model was found using two neural layers, the first layer containing ten Tanh neurons, and the second layer containing two Tanh neurons. The final model performed with an R^2 of 0.817, 0.812, and 0.567 on the training, validation, and test data, respectively. Notably, running this same model again did not produce good results (likely due to the low number of data points used), therefore it is plausible to assume that if more test data was collected, the model may not perform well against it. Fit performance plots such as the actual versus predicted, residuals, model fit error (MFE) and model representation error (MRE) for this model can be seen in Figure 9 below. Note that if the model predicted an adjusted TRL below 0, the adjusted TRL was defaulted to 0.358 (TRL 1), and if the model predicted an adjusted TRL above 9, the adjusted TRL was defaulted to 9 (TRL 9).

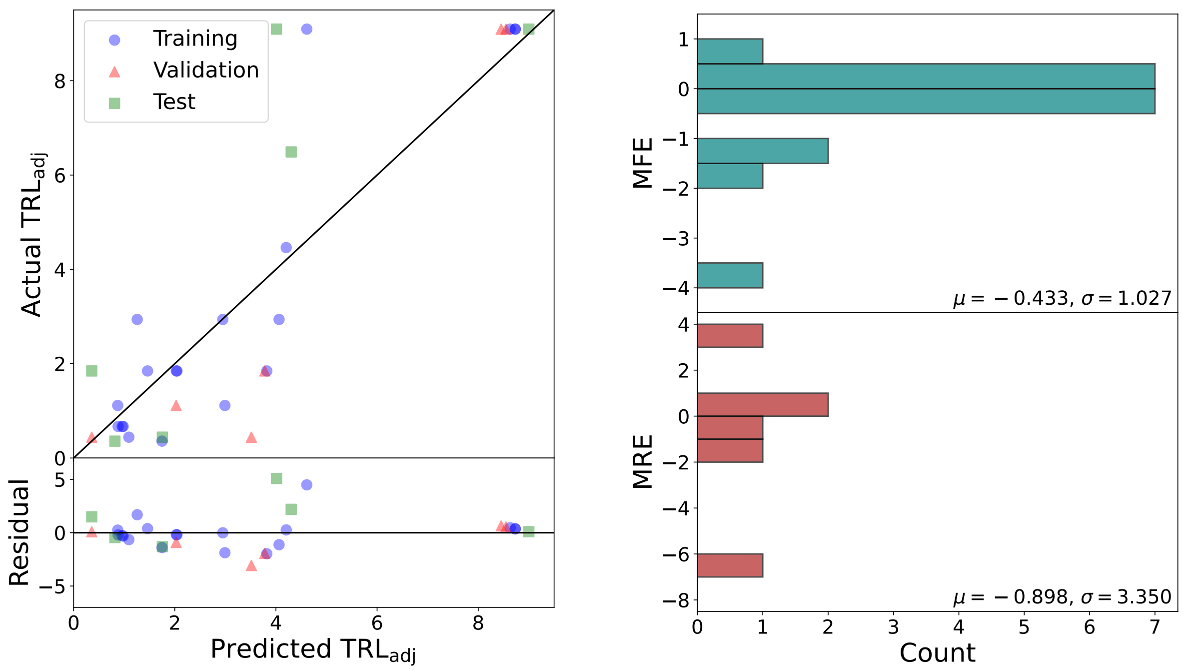


Fig. 9 Model performance metrics for the TRL adjusted prediction model. Left: actual versus predicted and residual versus predicted plots, where the black lines represent the ideal outcome. Right: MFE and MRE computed using the training and validation data, respectively (the mean, μ, and the standard deviation σ have also been shown).

From the left plot in Figure 9, one can observe that although the model does have some predicting power, it is over predicting TRL adjusted values between 3-5, which approximately corresponds to TRL values between 6-7. This is likely because of the initial data set only containing four technologies between TRL 6-7, corresponding to the growth phase of the TLC. Additionally, the number of patents over time was used as an indicator of technologies in this growth phase, however for all the technologies queried between TRL 6-7, no patents were found (as can be seen in Table 1 and 2). As such, the model is unable to extrapolate how technologies in this TRL range should behave with respect to the number of patents. Another possible cause for the poor overall prediction power was the S-curve fittings. Fitting S-curves to each time trace was found to sometimes be problematic as technologies that have few results per year may be fit poorly due to a lack of data. For example, only one NSF award was found for solar sail propulsion, yet SciPy fit an S-curve which started at zero and ended at one, so that the S-curve value indicated a fully mature technology. This clearly cannot be concluded given that only one data point was found. Methods to consider outliers such as this must be developed in future work.

Investigating further, one can turn to the MFE and MRE distributions to determine outliers and see what these outliers have in common. Looking at the bars located further away from the mean (around zero), these points all correspond to technologies in the emerging phase of the TLC (TRL 1-5), where most of the errors result from an over prediction in the adjusted TRL value. These technologies all had a significant number of Spinoff results, even though they most likely should not have commercial products available given their low maturity. This is likely caused by the inability to use logical expressions with the Spinoff API (as described in the previous section), leading to looser search queries being used. Consequentially, for future studies other APIs and databases should be considered for the mature phase to provide a better indication of a mature technology, and not provide false positives for immature technologies.

Even with these drawbacks, the model still manages to have a relatively high R^2 when predicting adjusted TRL values. Going back to the original task of predicting TRL, the adjusted values can then be converted to TRL by taking the inverse of Eq.2 and rounding the result to the nearest TRL value. The result of this ordinal regression problem can then be shown using a confusion matrix, which summarizes outcomes in a classification problem by visualizing actual versus predicted outcomes on a percentage basis using a heatmap. An ideal outcome has all the results along the diagonal (bottom left to top right) indicating that all predicted values exactly matched the actual outcomes. The confusion matrices for the training, validation, and test data can be seen below in Figure 10.

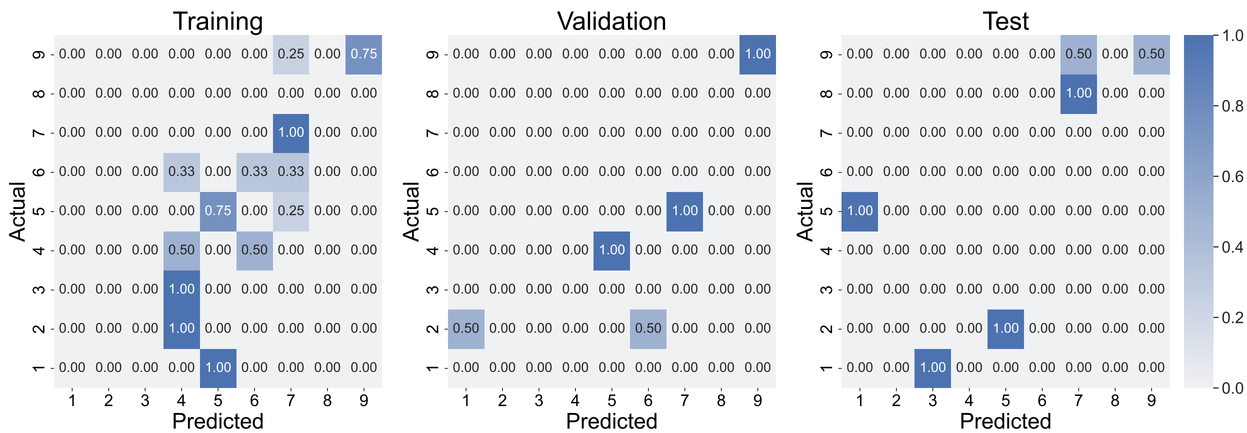


Fig. 10 Normalized confusion matrix for the training, validation, and test data showing the models predicted TRL classification against the actual classification (if all technologies were classified correctly, only boxes on the diagonal from the bottom left to top right would be shaded in).

These confusion matrices show that the model does not perform well when trying to classify low TRL values, as was previously determined when looking at the MFE and MRE results. Although, interestingly it seems as though the model is consistently predicting the same incorrect TRL value for these low TRL technologies. This may be improved by adding more technologies with low TRL values to the training data set to give the model more data to make better predictions. The model predicts close to the actual result for higher TRLs, often being off from the exact value by one or two on the training data, while significantly varying for the validation and test data. Again, this is likely caused by the low number of data points used to create this model (19 training points gives an average of 2 points per TRL level). This lack of data points per TRL level can be counteracted by instead of converting from TRL adjusted values to TRL, converting these adjusted values to TLC stages by grouping TRL levels together. The following confusion matrices are produced by grouping technologies between TRL 1-5 under emerging technologies, TRLs 6-7 under growing technologies, and TRLs 8-9 under mature technologies.

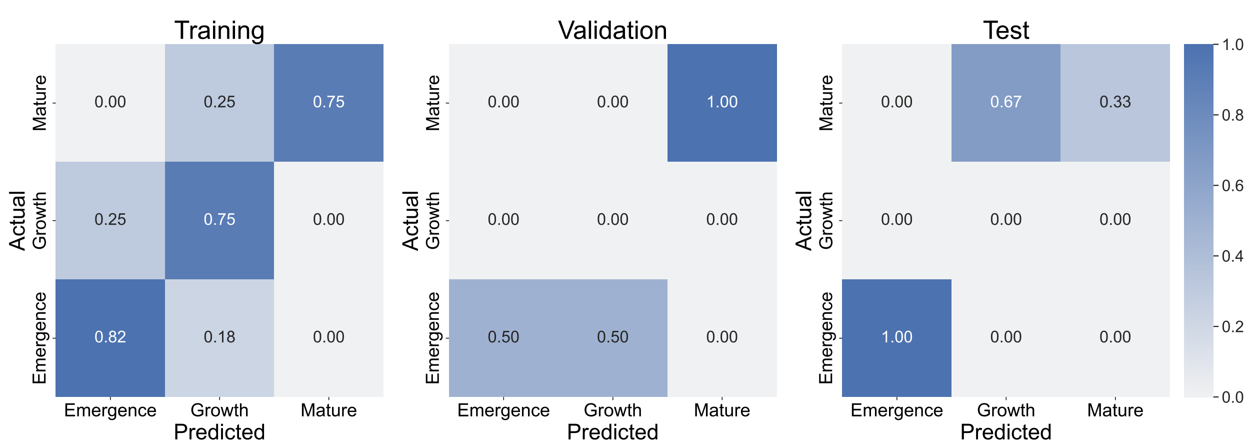


Fig. 11 Normalized confusion matrix for the training, validation, and test this time showing TRL predictions grouped into emergent (TRL 1-5), growth (TRL 6-7), and mature TLC phases (TRL 8-9).

The model performs much better for this classification problem, indicating that the chosen bibliometric factors do indeed indicate different stages in a technologies life cycle. Comparing to the previous results suggests that these factors alone may not be sufficient to then expand the classification task to the entire TRL scale. For this, more factors may be required that can be applied at each TRL level individually or for smaller groupings of TRL levels. Additional factors also allow for greater robustness in case few or no results can be found for another metric. Notably, as bibliometric factors are environmental by nature, the type of technology being considered likely impacts the metrics being collected. As such, it is hypothesized that a better bibliometric model can be created by considering only one type of technology. The suggested method would be to base the chosen technology type off of NASA's 2015 Technology Roadmap categories and to then select technologies within that category to build a model. This does limit the number of technologies available to be used to create the model, however it reduces uncertainty resulting from different technologies behaving differently.

Overall, the final model has a number of drawbacks associated with it that make it unlikely to be able to accurately predict TRL for other technologies. However, a substantial amount of insight and knowledge has been gained on how to build these models. For future work, it is suggested that authors take greater care in developing search queries, collect significantly more data (ensuring that each TRL level is fairly represented), consider a greater number and variety of bibliometric factors, attempt different ordinal regression techniques, and isolate the model to only apply to one type of technology. Bibliometric models are still a key enabler for the automation of TRL predictions which will then further enable the automation of mission budget and schedule forecasts, greatly reducing time spent at the conceptual phase determining whether a mission is feasible. The simplistic naive model shown here is already able to classify TLC phases at a sufficient degree of accuracy, therefore one can be optimistic that an improved bibliometric model will be able to classify individual TRL levels to a similar degree of accuracy.

V. Conclusion

The need to accurately assess technology maturity or TRL is constantly growing as interest in cislunar space grows, with government, academic, defense, and commercial stakeholders all playing a part in this growth. Here we have presented a method, attempted before with fewer bibliometric factors [10], that utilizes machine learning techniques and bibliometric data to automate the TRA process. It is shown that a simple machine learning regression model is able to predict TLC phases with sufficient accuracy using bibliometric data, however when trying to extend this to predict specific TRL levels, the model performs increasingly poorly on new test data largely due to the small number of data points used to generate the model. Prediction of the correct TLC phase is still incredibly useful as it characterizes a technologies maturity in a major development stage, allowing for the transition to more specific models that apply at each TLC phase, or for basic considerations of cost and schedule.

Future work should focus on exploring more bibliometric factors with more sophisticated ordinal regression techniques. This will allow for the creation of more accurate models, although these will require the curation of larger technology data sets to enable more complicated machine learning techniques to be used. Subject matter experts can also be brought in initially to determine search queries in order to further validate this bibliometric approach, and eventually develop methods for writing search queries for a general technology that can be used by anyone. Thus, removing the necessity for subject matter experts in the TRA process, and automating the budget and schedule forecasting process, allowing for more complicated missions to be conceived and tested for feasibility earlier in the development process.

Appendix

Table 1 Technologies selected from NASA's 2015 Technology Roadmap [20] to be used to generate a bibliometric TRL prediction model. Plain text search queries for each technology have also been shown.

Technology	TRL 2015	Search Query
	Training	
3D Imaging Sensor	7	"3D imaging sensor"
Acid-Resistant Solar Array Structures	1	"Acid resistant solar array"
Advanced Respirator	6	"Chemical rebreather" OR "Space respirator"
Advanced sensors for 3D terrain mapping	5	"Re-entry terrain mapping sensor" OR "3D terrain mapping sensors"
Autonomous aerobraking	5	"Autonomous aerobraking"
CO2 Removal (Closed-Loop)	2	"Space closed loop CO2 removal"
Deep Robotic Drilling	4	"Deep robotic Drilling"
Deep-space Position System	9	"Deep space positioning system" OR "DPS"
Dust Removal Brush	3	"Dust removal brush"
Free-water Shower for Full Body Cleansing	6	"Space shower" OR "Space bath"
Hybrid Propulsion	3	"Space hybrid propulsion"
Inflatable entry systems	5	"Inflatable entry system" OR "HIAD" OR "Hypersonic inflatable Aerodynamic Decelerator"
Integrated Docking and Automated Rendezvous Systems Design	3	"Autonomous rendezvous and docking"
LIDAR	9	"Light detection and ranging" OR "LIDAR"
Onboard Real-Time Planning and Scheduling	9	"Automated space planning and scheduling"
P50-Cork	9	"P50 Cork" OR "P50-Cork"
Regolith Sample Handling and Transfer	6	"Regolith handling and transfer" OR "Space rock handling and transfer"
Solar Sail Propulsion	5	"Solar sail propulsion"
Wheels for Planetary Surfaces	4	"Pneumatic terrestrial wheels" OR "Planetary wheels"
	Validation	
Beamed Energy Propulsion	2	"Beamed energy propulsion"
Dust Covers	4	"Dust cover" OR "Dust intrusion coating"
Entry Guidance Software	5	"Entry guidance software"
Lunar Habitat	2	"Lunar habitat"
Photovoltaic Solar Cell	9	"Solar cell"
PICA	9	"PICA" OR "Phenolic Impregnated Carbon Ablator"
	Test	
Cabin Air Sensor	8	"Space cabin air sensor"
Lightweight Crew Quarters	2	"Lightweight crew quarters"
RS-25 Engine	9	"RS-25 Engine"
Stabilized Foods	9	"Space food packaging"
Transformable or Morphable Entry Systems	1	"Transformable entry system" OR "Morphable entry system"
Wastewater Collection	5	"Wastewater storage" OR "Wastewater processing"

Table 2 API and S-curve fit results for the technologies and queries shown in Table 1. TRL_{adj} has been calculated using Eq.2. NR means that no results were found.

Technology	S-curve Science	S-curve NSF	S-curve Patents	Spinoff	Spinoff Norm	TRL_{adj}
			Training			
3D Imaging Sensor	0.708	0.658	NR	257	0.057	4.462
Acid-Resistant Solar Array Structures	1.000	NR	NR	136	27.200	0.358
Advanced Respirator	0.947	NR	NR	53	4.818	2.938
Advanced sensors for 3D terrain mapping	NR	NR	NR	64	NR	1.846
Autonomous aerobraking	0.714	NR	NR	14	0.452	1.846
CO2 Removal (Closed-Loop)	0.687	NR	NR	37	0.725	0.442
Deep Robotic Drilling	0.864	NR	NR	88	1.872	1.114
Deep-space Position System	0.818	0.893	0.942	45	0.157	9.094
Dust Removal Brush	0.997	1.000	NR	32	2.133	0.67
Free-water Shower for Full Body Cleansing	0.746	NR	NR	114	0.072	2.938
Hybrid Propulsion	0.560	NR	NR	276	0.462	0.67
Inflatable entry systems	0.978	NR	1.000	12	0.188	1.846
Integrated Docking and Automated Rendezvous Systems Design	0.837	NR	0.902	72	0.242	0.67
LIDAR	0.490	NR	0.709	98	0.018	9.094
Onboard Real-Time Planning and Scheduling	0.820	NR	NR	102	0.685	9.094
P50-Cork	NR	NR	NR	2	0.500	9.094
Regolith Sample Handling and Transfer	1.000	NR	NR	83	83.000	2.938
Solar Sail Propulsion	0.947	NR	1.000	226	0.381	1.846
Wheels for Planetary Surfaces	0.999	0.870	NR	25	1.042	1.114
			Validation			
Beamed Energy Propulsion	0.897	NR	NR	295	3.172	0.442
Dust Covers	NR	0.970	NR	58	0.247	1.114
Entry Guidance Software	0.782	NR	NR	248	1.953	1.846
Lunar Habitat	0.846	NR	0.981	32	0.041	0.442
Photovoltaic Solar Cell	0.087	0.333	0.882	135	0.001	9.094
PICA	0.961	0.999	0.798	26	0.268	9.094
			Test			
Cabin Air Sensor	0.544	NR	NR	246	5.125	6.49
Lightweight Crew Quarters	NR	NR	NR	101	25.250	0.442
RS-25 Engine	1.000	0.880	NR	125	0.001	9.094
Stabilized Foods	0.811	NR	NR	179	0.772	9.094
Transformable or Morphable Entry Systems	NR	NR	NR	105	15.000	0.358
Wastewater Collection	0.703	0.964	0.985	79	0.012	1.846

Acknowledgments

This research is supported by the Georgia Tech Cislunar Architecting Initiative. The authors would like to thank Leon Chen and Angelly Castro of the School of Aerospace Engineering at the Georgia Institute of Technology for their thoughts and input on this work. Lastly, the authors would like to thank the Aerospace Systems Design Laboratory for its support on this work, with special thanks to the cislunar architecture team consisting of: Dr. Bradford Robertson, Dr. Selcuk Cimtalay, Daewoon Kim, Leon Chen, Stephanie Zhu, Jacob Zhong, Timothy Elrick, Jacob Hawkins, Kenneth Smith, Johann Bradley, Paul Wang, Angelly Castro, Tabitha D'amato, Sophia DeSisto, Alfonso Lagares de Toledo, and Trevor Banks.

References

[1] GAO, "NASA: assessments of major projects 2021," , May 2021. URL `https://www.gao.gov/assets/gao-21-306.pdf`.

[2] Senate, "Programs in Peril: An Overview of the GAO High Risk List," , February 2005. URL `https://www.govinfo.gov/content/pkg/CHRG-109shrg20174/html/CHRG-109shrg20174.htm`.

[3] Martino, J. P., "A review of selected recent advances in technological forecasting," *Technological Forecasting and Social Change*, Vol. 70, No. 8, 2003, pp. 719–733. https://doi.org/10.1016/S0040-1625(02)00375-X.

[4] GAO, "Technology Readiness Assessment Guide," *GAO Reports GAO-16-410G*, 2016.

[5] Young, P., "Technological growth curves: A competition of forecasting models," *Technological Forecasting and Social Change*, Vol. 44, No. 4, 1993, pp. 375–389. https://doi.org/10.1016/0040-1625(93)90042-6.

[6] Debecker, A., and Modis, T., "Determination of the uncertainties in S-curve logistic fits," *Technological Forecasting and Social Change*, Vol. 46, No. 2, 1994, pp. 153–173. https://doi.org/10.1016/0040-1625(94)90023-X.

[7] Watts, R. J., and Porter, A. L., "Innovation forecasting," *Technological Forecasting and Social Change*, Vol. 56, No. 1, 1997, pp. 25–47. https://doi.org/10.1016/S0040-1625(97)00050-4.

[8] Lezama-Nicolás, R., Rodríguez-Salvador, M., Río-Belver, R., and Bildosola, I., "A bibliometric method for assessing technological maturity: the case of additive manufacturing," *Scientometrics*, Vol. 117, No. 3, 2018, pp. 1425–1452.

[9] Faidi, S., "Assessing Bibliometrics for the Automation of Technology Readiness Level Assessments," Master's thesis, 2021.

[10] Cauthen, K., Rai, P., Hale, N., Freeman, L., and Ray, J., "Detecting technological maturity from bibliometric patterns," *Expert Systems with Applications*, Vol. 201, 2022, p. 117177. https://doi.org/10.1016/j.eswa.2022.117177.

[11] Mankins, J. C., "Technology readiness levels," *White Paper*, Vol. 6, 1995.

[12] Mankins, J. C., "Technology readiness assessments: A retrospective," *Acta Astronautica*, Vol. 65, No. 9, 2009, pp. 1216–1223. https://doi.org/10.1016/j.actaastro.2009.03.058.

[13] Malone, P. K., "Applying system readiness levels to cost estimates — A case study," *2018 IEEE Aerospace Conference*, 2018, pp. 1–18. https://doi.org/10.1109/AERO.2018.8396455.

[14] Nieto, M., Lopéz, F., and Cruz, F., "Performance analysis of technology using the S curve model: the case of digital signal processing (DSP) technologies," *Technovation*, Vol. 18, No. 6, 1998, pp. 439–457. https://doi.org/10.1016/S0166-4972(98)00021-2.

[15] Terrile, R. J., Doumani, F. G., Ho, G. Y., and Jackson, B. L., "Calibrating the Technology Readiness Level (TRL) scale using NASA mission data," *2015 IEEE Aerospace Conference*, 2015, pp. 1–9. https://doi.org/10.1109/AERO.2015.7119313.

[16] Tang, D., Wiseman, E., and Archambeault, J., "Using Four-Quadrant Charts for Two Technology Forecasting Indicators: Technology Readiness Levels and R&D Momentum," 2012.

[17] Olechowski, A., Eppinger, S. D., and Joglekar, N., "Technology readiness levels at 40: A study of state-of-the-art use, challenges, and opportunities," *2015 Portland International Conference on Management of Engineering and Technology (PICMET)*, 2015, pp. 2084–2094. https://doi.org/10.1109/PICMET.2015.7273196.

[18] Gil, M., Wróbel, K., Montewka, J., and Goerlandt, F., "A bibliometric analysis and systematic review of shipboard Decision Support Systems for accident prevention," *Safety Science*, Vol. 128, 2020, p. 104717. https://doi.org/10.1016/j.ssci.2020.104717.

[19] Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and SciPy 1.0 Contributors, "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nature Methods*, Vol. 17, 2020, pp. 261–272. https://doi.org/10.1038/s41592-019-0686-2.

[20] Chandler, F., Ambrose, R., Biegel, B., Brown, T., Carter, J., Culbert, C., Edwards, C., Fox, J., Glaessgen, E., Hurlbert, K., et al., "NASA Technology Roadmaps," *NASA*, 2015.

[21] Singh, V. K., Singh, P., Karmakar, M., Leta, J., and Mayr, P., "The journal coverage of Web of Science, Scopus and Dimensions: A comparative analysis," *Scientometrics*, Vol. 126, No. 6, 2021, pp. 5113–5142.

[22] Cozzens, S., Gatchair, S., Kang, J., Kim, K.-S., Lee, H. J., Ordóñez, G., and Porter, A., "Emerging technologies: quantitative identification and measurement," *Technology Analysis & Strategic Management*, Vol. 22, No. 3, 2010, pp. 361–376. https://doi.org/10.1080/09537321003647396.

[23] The National Science Foundation, "NSF Awards Database," , 2022. retrieved 18 Oct 2022, https://www.nsf.gov/awardsearch/.

[24] PatentsView, "PatentsView Database," , 2022. retrieved 23 Oct 2022, https://patentsview.org/.

[25] NASA Spinoff, "Spinoff Database," , 2022. retrieved 23 Nov 2022, https://spinoff.nasa.gov/.

[26] Gutiérrez, P. A., Perez-Ortiz, M., Sanchez-Monedero, J., Fernandez-Navarro, F., and Hervas-Martinez, C., "Ordinal regression methods: survey and experimental study," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 28, No. 1, 2015, pp. 127–146.

[27] Conrow, E. H., "Estimating Technology Readiness Level Coefficients," *Journal of Spacecraft and Rockets*, Vol. 48, No. 1, 2011, pp. 146–152. https://doi.org/10.2514/1.46753.