Exploring the Applicability of Machine Learning to Roman Egypt

A New Facet of Ancient Demography

Peter Richards

Biohistory of the Ancient Mediterranean

Professor Nathan L. Pilkington, Ph.D.

May 5th, 2017

Abstract

Classical historical demography has long relied on traditional statistics for insight. As this path is well-trod, new insight is difficult to glean with traditional methods. This paper aims to show that the burgeoning field of machine learning offers an exciting opportunity to revisit existing data with greater powers of prediction and superior proxies.

Introduction

The great challenge for demographers of the ancient Roman world is a lack of sources upon which to base their study. A smattering of literary sources, gravestones, epitaphs, archaeological survey evidence, bioarchaeological evidence, paleodemographic evidence, and the census returns of Roman Egypt are the only sources for ancient Roman demographers to draw upon.[1] Of these sources, the census returns offer unique promise. Free from the biases of gravestones and literature, and less dependent on speculation than archaeological survey, bioarchaeological, or paleodemographic evidence, the census returns offer an unparalleled window into the demography of the ancient Roman world. [2] This unique opportunity means the methodology by which the returns are studied is takes on an outsized importance.

Traditionally, ancient historians have relied on statistics for their demography. Tried and true, traditional statistics offers the predictive ability necessary for extrapolating the composition of ancient populations. However, with the advent of the computing age, traditional statistics is no longer always the best option for demography. A new method, machine learning, is capable of

[1] Hin, Saskia, "*Ancient Demography*" (Oxford Bibliographies, 2015), 1
[2] Keith Hopkins wrote in his 1980 article on brother-sister marriage in Roman Egypt that "[the census returns] are far from perfect, but … the best data we have." Beggars can't be choosers when dealing with 2,000 year old data.

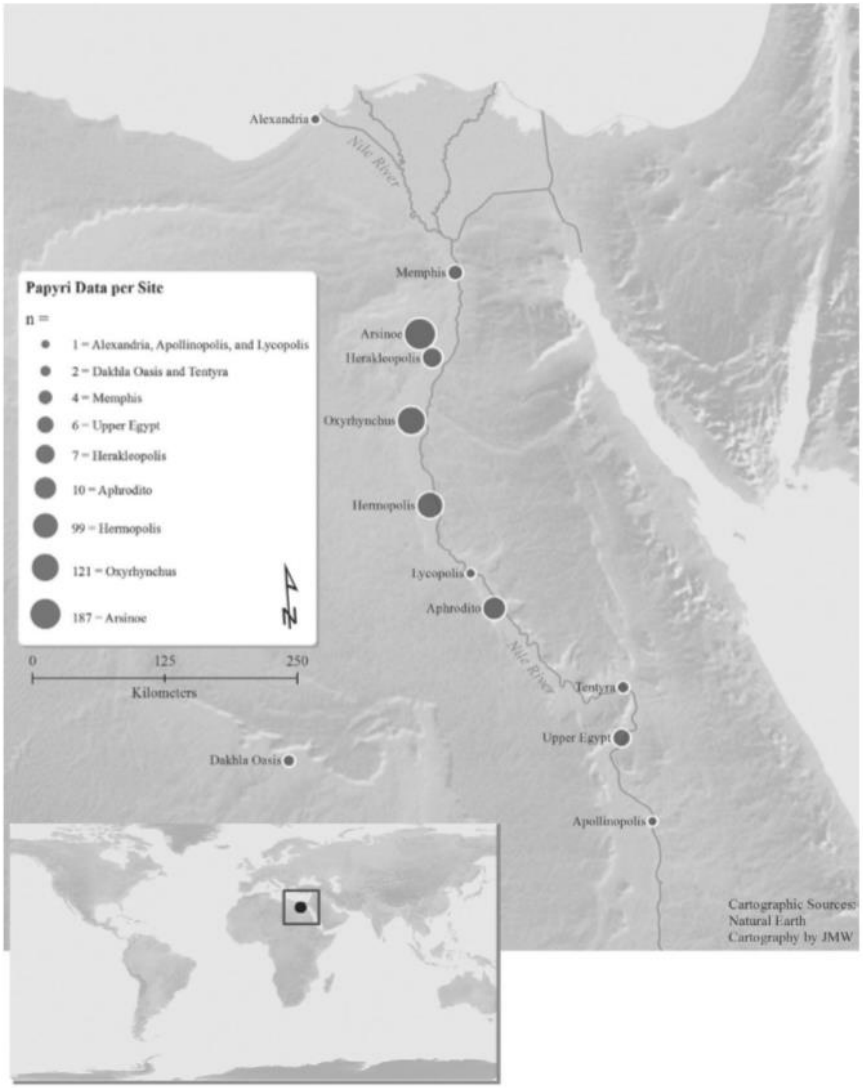
making superior predictions to those of traditional statistics under the right conditions. Machine learning offers a novel perspective and innovative techniques for dealing with the challenges of the census returns, and its adoption by classical demographers will allow original insight.

Roman Census Returns

The returns themselves are a collection just over three hundred census declarations filed by ordinary Egyptians. They date from 11/12 AD to 257/258 AD, with a high concentration in the second and early third centuries. [3] Past the irregularities of the early to mid-first century, the census occurred every fourteen years, presumably because fourteen was the age at which boys became subject to poll tax. [4] Most of the returns come from middle Egypt, from cities and villages along the Nile, and approximately 92% of them come from the nomes of Arsinoite, Oxyrhynchite, and Hermopolite.[5]

---

[3] Bagnall, Roger S., and Bruce W. Frier, "*The demography of Roman Egypt*" (Cambridge, 1994), xv
[4] Ibid., 27
[5] Harper, Kyle, "*People, Plagues, and Prices in the Roman World: The Evidence from Egypt*" (*The Journal of Economic History*, Vol 76, No. 3, 2016), 810



[6]

At their core, the returns are essentially a tool of state, and there are many theories explaining their role. Because no imperial order, prefectural edict, or internal correspondence of the Roman government exists outlining the purpose of the census, all of these theories are educated speculation.[7] In all likelihood the census was carried out for a mix of reasons, as the amount of data gathered does not fit neatly into any one purpose. Historians Roger Bagnall and

[6] Harper (2016), 811
[7] Bagnall and Frier (1994), 27

Bruce Frier suggest two broad reasons, taxation and control of the population, and speculate further that the census was an effective means of projecting Roman power.[8]

For demographic analysis however, the purpose of the census is less important than what it contains. Most returns contain a head of household, his dependents, his location, and the year the census was taken. Unfortunately, this is the ideal and not the rule, and the challenges of the data set itself are many. First, the data points are spread out over hundreds of years and hundreds of miles, making coherent statistical analysis difficult. Second, many data points are incomplete, lacking place, year, or a full roster of who the census was purportedly counting. Such is to be expected from the fragments thousands of years old. Third, interpreting the data set relies on imperfect[9] and often contentious translating, resulting in additional uncertainty. And last but not least, the data set is relatively small for what other historians have attempted to accomplish with it. As Scheidel writes, the census returns represent at best 1/80,000 of the total population of Egypt at the time, making coherent statistical analysis very difficult.[10] Trying to determine the demography of Roman Egypt from the census returns is similar to trying to determine the population of Wisconsin without any data from Milwaukee, and just a smattering of a few hundred families concentrated in the north of the state, mostly in Wausau.

In light of all of these limitations, insightful analysis of the census returns is quite difficult. With such a small data set, statistically significant results can be hard to come by

[8] Bagnall and Frier (1994), 28-30

[9] Bagnall owns up to this in the appendix of The Demography of Roman Egypt, and writes “I am under no illusion that the texts as corrected are perfect” (Bagnall and Frier (1994), 180). When dealing with 2,000 year old papyrus, translation with absolute certainly is unfortunately not universally achievable. And because of how little data we have, even partial returns are valuable data points.

[10] Scheidel, Walter, “*Death on the Nile: Disease and the Demography of Roman Egypt*” (Leiden, 2001), 51

without large assumptions. That's not to say previous historians haven't tried however, and to see the power and advantage of machine learning over traditional methods of ancient demography, first it is helpful to look at what came before.

Notable Previous Scholarship

Bagnall and Frier's book *The Demography of Roman Egypt* is a natural starting point for looking at previous scholarship, as their work in bringing the census returns together really put them on the map. To carry out their statistical analysis, Bagnall and Frier relied heavily on model life tables. From those tables they drew conclusions about the composition of ancient Egyptian households, male and female life expectancy, sex ratios, and age distribution. After this basic statistical analysis, Bagnall and Frier applied their findings to speculate about marriage, female fertility, extramarital birth, and slave birth. While a valiant effort, the book contains a few fatal flaws, most connected to the aforementioned use of model life tables. Because making life tables is such an intensive process, there isn't much of a selection, and the nature of statistical modeling is such that Bagnall and Frier can coerce the already small data set to fit cleanly into the model life table they need. Unfortunately, the ability to make the data fit invalidates many of their conclusions.

After Bagnall and Frier corralled and put the census returns on the map, others employed them for larger demography projects, such as Lo Cascio in his study of the population of the Roman Empire as a whole[11] and Kyle Harper in his larger study of the Roman Economy. [12] No

[11] Lo Cascio, Elio, "*The Size of the Roman Population: Beloch and the Meaning of the Augustan Census Figures*" (The *Journal of Roman Studies*, Vol. 84, 1994), 35
[12] Harper (2016), 806

one digs into them more than Walter Scheidel however. Scheidel's book *Death on the Nile: Disease and the Demography of Roman Egypt* is the next milestone in historical scholarship regarding the returns because of the breadth it is able to cover. With a potent combination of the census returns and epitaphs from ancient and mediaeval Egypt and Nubia, Scheidel explores mortality patterns, causes of death, age structure, life expectancy, and even population size and demographic change. In doing so he levels a withering critique of model life tables, exposing their flaws and inherent biases. Scheidel's book is not without flaw either however. He settles for a wishy-washy conclusion, that we must be content to read the census returns from Egypt as evidence for merely "diversity at a local level, emphasizing the often intractable complexity of demographic conditions in the past."[13] Clearly afraid to say anything too definitive lest it be attacked with the same fever he levies at those who came before him, Scheidel is content with wide ranges and pages and pages of assumption justification.

The book is the logical culmination of traditional statistical modeling, a well-researched analysis with, as previously stated, disappointingly wide ranges. Scheidel uses his conclusions to branch out into a number of other studies, often incorporating other evidence as well. He references the census returns in papers about Real Wage[14] and the Population Size of Rome[15], to Greco-Roman sex ratios[16] and the Antonine Plague[17]. Scheidel recognizes as well as anyone what a seismic shift proper interpretation of these records would have on our understanding of

13 Scheidel (2001), xxviii
14 Scheidel, Walter, "*Real Wages in Early Economies: Evidence for Living Standards from 1800 BCE to 1300 CE*" (*Journal of the Economic and Social History of the Orient*, Vol. 53, No. 3, 2010), 436
15 Scheidel, Walter, "*Roman population size: The logic of the debate*" (*Mnemosyne Supplements*, 2008), 17 - 70
16 Scheidel, Walter, "*Greco-Roman sex ratios and femicide in comparative perspective*" (*Princeton/Stanford Working Papers in Classics*, 2010a), 2
17 Scheidel, Walter, "*Roman wellbeing and the economic consequences of the Antonine Plague*" (*Princeton/Stanford Working Papers in Classics*, 2010b), 2-3, 6-7

the classical world, but he wisely won't commit to specific interpretation because of the limitations of traditional statistical modeling. The future then is not in the oversaturated area of traditional statistical modeling, well-trod ground that Scheidel has expertly passed over more than a half dozen times, but in machine learning.

Machine Learning: Overview & Applicability

So what is machine learning? And how is it applicable? Machine learning is "the process of discovering the relationships between predictor and response variables using computer-based statistical approaches."[18] In other words, machine learning takes known data, searches for patterns, and uses the patterns it observes to make predictions. A key part of this is process is improvement over time. The predictions are in the beginning very poor, but as more data is processed they become much more accurate. While machine learning is excellent at distilling millions and sometimes billions of data points into one coherent prediction, it can struggle with too little data. Low data machine learning is an area of ongoing research, but initial findings are promising for making insights not possible with traditional statistics. Because machine learning is automated once the code is written, the model merely needs to be given a human generated data set to make a prediction, so historians will be able to create many different potential models based on different historical proxies. This will allow an avoidance of the model life table problem previously discussed, where scholarship must revolve around one imperfect proxy because of how much time and effort it takes to create. This combination of lower effort models and new insight into the same data will allow machine learning to contribute to our understanding of the census records. As long as the risks and drawbacks are properly understood,

[18] Witten, I.H. and E. Frank "*Data Mining: Practical Machine Learning Tools and Techniques*" (San Francisco, 2005), 215

machine learning will be a powerful tool for the ancient demographer, able to avoid the pitfalls of lifetables and create far superior models.

Interestingly enough, traditional statistics and machine learning both aim to do the same thing, and as machine learning advances the lines between the fields are beginning to blur. Historically, tension existed between the two communities,[19] and while it has simmered down, there are key distinctions to keep in mind. The biggest difference is statistics starts with a model, for as a subgenre of mathematics it is very interested in laying out all of its assumptions and methods for critical analysis from the beginning. Machine learning builds its model automatically, from existing data, automating the process and removing the human element of modeling building. There is a fluidity to this process, and an ability to adapt to the data as it is interpreted as a means of increasing accuracy. This allows modeling at previously impossible scales, and allows a potentially superior model, given the right data. Note that both techniques are beholden to the data they are given however, and the predictions' accuracy is heavily reliant on the accuracy of the data it is given.

In light of these differences, it is becoming clearer why machine learning, and not traditional statistics, is a better choice for analyzing Egyptian census records. A modern proxy, a recent paper published by the NIH exploring potential uses of machine learning for future demographic predictions, helps to illustrate this advantage.[20] Seeing the problem with the use of population surveys to track regional public health statistics, because these surveys are high in

[19] Breiman, Leo, "*Statistical Modeling: The Two Cultures*" (*Statistical Science*, Vol. 16, No. 3, 2001), 199-215
[20] Luo et al., "*Is Demography Destiny? Application of Machine Learning Techniques to Accurately Predict Population Health Outcomes from a Minimal Demographic Dataset*" (*PLoS One*, 2015), 1-13

cost and slow to report their data, researchers in NIH tested the applicability of machine learning to solve the problem. Now instead of waiting years for data on the health of state populations, researchers wanted to build a tool capable of “inferring regional health outcomes from socio-demographic data.”[21] The tool was successful. After feeding the data of 30 states into their model, they were effectively able to predict with statistical significance the data of the remaining 20 states, and compare that data to what was observed. They found their predictions “were highly correlated with the observed data, in both the states included in the derivation model (median correlation 0.88) and those excluded from the development for use as a completely separated validation sample (median correlation 0.85), demonstrating that the model had sufficient external validity to make good predictions, based on demographics alone, for areas not included in the model development.”[22] The implications for classical demography are immense. Machine learning is capable of producing models predicting the health of a population with the right techniques and enough health data about similar populations.

This proxy does a good job of illustrating the applicability of machine learning to demography, and is a wonderful proof of concept, but unfortunately census records we are concerned with are a special case. Unlike in the above example, we know far less about the existing demography of the Egyptian population and machine learning does best with as many data points as possible, so the model will need to be built with proxies. For this reason, their methodology is not directly applicable, although some of its techniques are. As the field is relativity new and constantly evolving, this is an area of current research, but as techniques develop classical demographers will find more and more for them in the field of machine

[21] Luo et al. (2015), 1
[22] Ibid., 2

learning. To demonstrate what could potentially be done, two recent examples stand out. They are notable not for their work with demography, but for their application of machine learning techniques to data sets with similar characteristics to the census records, namely the small size.

First Small Data Set Machine Learning Technique: Synthetic Noise

The first proxy for dealing with small data discusses the technique of "synthetic noise." This technique is very recent, and was developed in the last year. Its creator, Igor Shuryak, set out to address some of the problems with studying the effects of radioactive contamination, namely that because of the danger associated with radioactivity most gathered data sets are small. Traditional statistical analysis with generalized linear models (GLMs) is underwhelming, because of a low amount of data points. To overcome this issue, Shuryak proposes some groundbreaking new techniques:

> We propose that analysis of small radioecological data sets by GLMs and/or machine learning can be made more informative by using the following techniques: (1) adding synthetic noise variables to provide benchmarks for distinguishing the performances of valuable predictors from irrelevant ones; (2) adding noise directly to the predictors and/or to the outcome to test the robustness of analysis results against random data fluctuations; (3) adding artificial effects to selected predictors to test the sensitivity of the analysis methods in detecting predictor effects; (4) running a selected machine learning method multiple times (with different random-number seeds) to test the robustness of the detected "signal"; (5) using several machine learning methods to test the "signal's" sensitivity to differences in analysis techniques.[23]

He applied these techniques to two case studies, one regarding the "bacterial abundance in soil samples under a ruptured nuclear waste storage tank" and the other "fungal taxa in samples of soil contaminated by the Chernobyl nuclear power plan accident." In both cases the machine

[23] Shuryak, Igor, "*Advantages of Synthetic Noise and Machine Learning for Analyzing Radioecological Data Sets*" (PLoS One, 2017), 1-19

learning techniques utilized allowed the researchers to extract useful information from the data that traditional statistical analysis could not see. Shuryak could create a graph able to differentiate signal and noise impossible with traditional statistical methods.

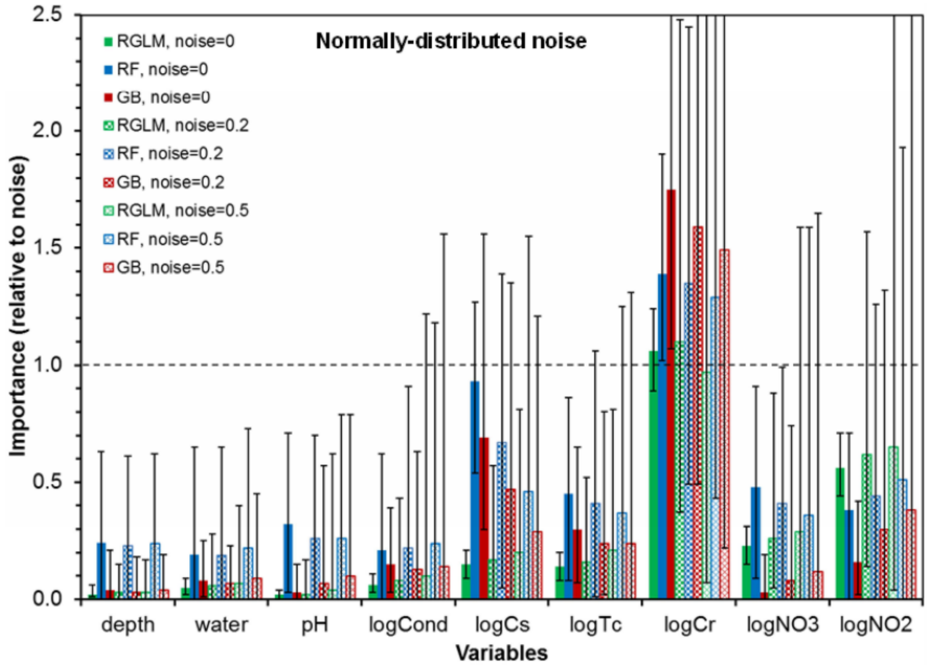


Fig 3. Summary of machine learning analyses of data set II (bacteria at Hanford) by the RGLM, RF and GB methods. The x-axis lists predictor variables, and the y-axis displays VIMr (bars = mean values for 1000 runs with different random number seeds, error bars = range). VIMr values > 1 indicate that the particular predictor achieved higher importance than any of the synthetic noise variables. The analyses were repeated under different amounts of normally-distributed noise (v), as listed in the legend. Details are described in the main text.

doi:10.1371/journal.pone.0170007.g003

[24]

The radioactive samples taken from both data sets take on new meaning after the analysis, and new ways of looking at the noise are possible. The study concludes “In summary, the proposed approach allows useful information to be extracted even from small and limited-quality radioecological data sets. Specifically, this approach was advantageous compared with the methodology used by the authors of references because it identified important patterns (i.e. the

[24] Shuryak (2017), 13

effect of radioactive contamination in data set I and the effect of Cr in data set II) which were not reported in the original publications."[25] While not a perfect analog, in the same way demographic information can be very complex and interdependent, so too can radioactive soil. In this instance, machine learning allows the influx of artificial data points designed to reveal statistical significance overlooked by traditional statistical methods, with great success. This technique is not unique to soil however, and there is no reason it couldn't succeed in the arena of ancient demography as well.

Second Small Data Set Machine Learning Technique: Mega Diffusion

Synthetic noise is not the only small data set machine learning technique however. Many older examples exist, most notably one called "mega diffusion." Developed by Li et al., the technique allows learning accuracy to be improved with statistical significance when applied to a very small data set.[26] The technique was developed in the context of flexible manufacturing system (FMS). FMS is "a manufacturing system that is not dedicated to the production of a single component, as in mass production, but is capable of producing a variety of components through computer control of a series of linked machine tools. FMS is particularly useful for the production of batches in relatively small numbers, say 50 to 500."[27] The key word of FMS is flexibility, which inevitably results in unpredictability. Before a company makes 500 of a new component, they want to make a few to make sure nothing is wrong. This leads to a small sample size however, and increasing the sample size is inconvenient because it costs time and money.

[25] Shuryak (2017), 18
[26] Li et al., "*Using mega-trend-diffusion and artificial samples in small data set learning for early flexible manufacturing system scheduling knowledge*" (*Computers & Operations Research*, Volume 34, Issue 4, 2007), 966-982
[27] Atkins, Tony, and Marcel Escudier, "A *Dictionary of Mechanical Engineering*" (Oxford, 2013), 26

Companies who offer FMS services are interested in learning all they can from that small sample size, and machine learning offers a great opportunity. Most standard machine learning techniques are not effective for this pursuit, so Li et al. set out to develop their own.

First, they outline numerous proposed solutions by other scientists, and show that their synthesis takes the best of all of them, creating a machine learning solution that works better than anything else for small data sets. It works by taking the advantages of the well-established Back Propagation Neural Network technique (BPNN) and utilizing it with far less data, using mega diffusion. Back Propagation Neural Network has been around since the late 1960s, and is the most widely applied neural network architecture.[28] Mega diffusion utilizes a few different mathematical formula to create additional data points for statistical analysis from what is already known, as pictured by this diagram.

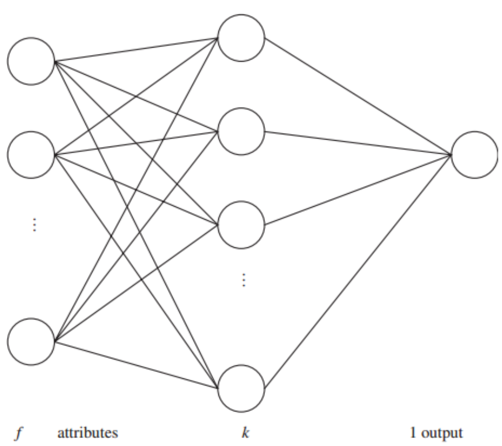


Fig. 4. The traditional $f - k - 1$ BPNN.

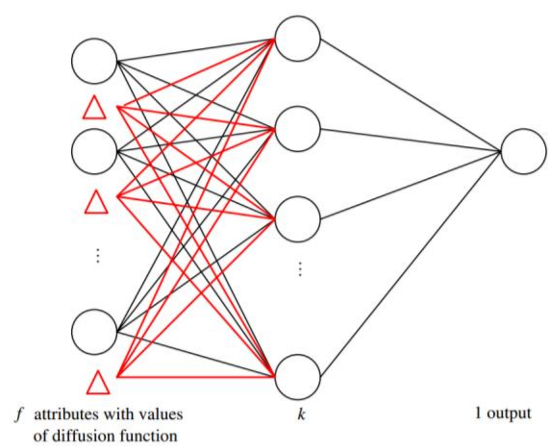


Fig. 5. The $2f - k - 1$ BPNN for small data set learning. [29]

[28] Hecht-Nielsen, Robert "*Theory of the Backpropagation Neural Network*" (*Neural Networks*, Volume 1, Supplement 1, 1988), 445
[29] Li et al (2007), 978

After the generation of these additional data points, BPNN therefore has more to run on, and as a result can draw more accurate conclusions. So how exactly are these data points generated? In short, the use of a diffusion function, which allows avoiding a common BPNN pitfall of training a BPNN directly from random (uniformly generated) samples.[30] The results of this technique is below.

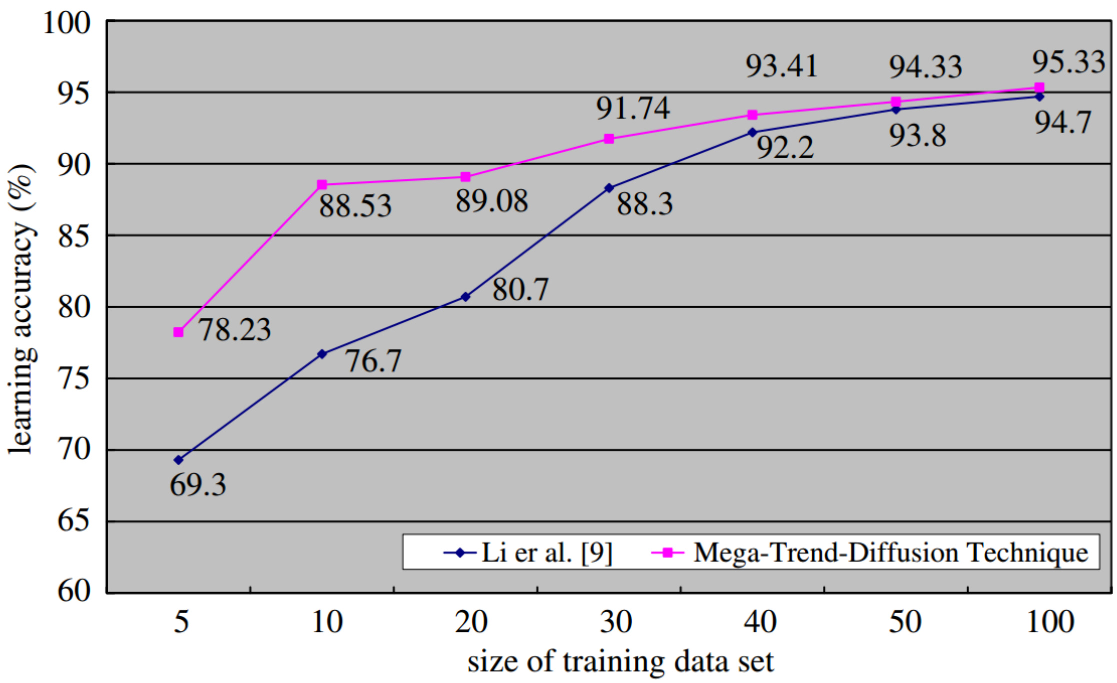


Fig. 6. The comparisons of computational results.

31

As shown by the above graph, the mega diffusion allows significantly superior learning accuracy with smaller data sets. BPNN does better with the additional data, even if it is artificially generated. This ability is again very relevant considering the small size of the census returns, as

[30] Ibid., 977

[31] Li et al (2007), 981

artificially generated data points capable of increasing machine learning accuracy would be quite helpful for understanding the census returns. Thinking about the census returns as a kind of neural network is also an interesting possibility, and fits in well with this machine learning method.

Machine Learning: Looking Forward

Unfortunately at this time, there is no direct proxy for a data set with both a small amount of variables and also a focus on demography. This is likely for two reasons. First, machine learning is in its infancy and as a result is cutting its teeth on easier cases with lots of data. Second, as a field machine learning has traditionally worked on modern day demography issues with lots of data, and not concerned itself with the past and its inevitable small data sets. Neither of these impediments are insurmountable however, and it is only a matter of time before the field matures and even better techniques develop. Prediction can be applied just as equally to the past as to the future.

Every day advancements are made in the field, from new techniques for smoothing to ideas about applicability to demography.[32] As the lines blur between traditional statistics and machine learning, data mining and artificial intelligence, new and exciting analysis will surely become possible. It is not dusk for ancient demography, as Scheidel and others have implied, it is dawn, for as computers get more powerful and newer techniques are developed the extrapolation

[32] Letouze, Emmanuel, "*Demography, meet Big Data; Big Data, meet Demography: Reflections on the Data-Rich Future of Population Science*" (New York, 2015), 1-44

that statistics and machine learning offer will only become easier and more accurate. Synthetic Noise and Mega Diffusion are the tip of the iceberg.[33]

Potential Historical Proxies

As the revolution comes however, we must circle back to the question of proxy. As previously stated, machine learning and statistics are only as good as the data they are fed. Returning to our first demographic example, the study found so much success because data from 30 states from 2007-2012 is applicable for understanding data from the 20 remaining states in 2007-2012. The Egyptian census returns lack such a clear proxy. The obvious impediment is the proxy must be unencumbered by modern medicine and hygiene, and data without these biases is few and far between. Our best bet is some combination of previously suggested ideas, most gathered by Scheidel in *Death on the Nile*. Scheidel uses a multitude of sources for his proxies, from Greek and Coptic funerary inscriptions to modern estimates of the population of Egypt. Machine learning does best with as direct an analog to what is being studied as possible, and two potential proxies stand out.

The first is the nineteenth century censuses of Egypt, whose categories of sex, kinship relation, age, and, legal status[34] among others match up remarkably well to the ancient census records. Using these records requires a tradeoff, as while their accuracy purportedly increased

[33] There is one crucial drawback that will need to be accounted for, one of the main reasons this sort of statistical analysis requires both an expert in machine learning and an expert ancient historian. Every single one of these experiments uses **random** subset of the data to extrapolate about the set as a whole, and the Egyptian census data is not random. Correcting for that bias will require mindfulness by all involved. It is impossible to speak more about this challenge without more knowledge of the eventual model that will be employed, but I do not think this bias is insurmountable.

[34] Dourmani, Beshara, "*Family History in the Middle East: Household, Property, and Gender*" (New York, 2003), 25-26

over time, by the late nineteenth century modern medicine and hygiene was prevalent enough to distort the direct proxy.[35] An additional layer of complexity also exists, because the modern European censuses did not break the population down to a house by house level, meaning acting as a direct proxy to the Roman census is impossible. None of these censuses is perfect, so again it will be up to the discerning ancient demographer to select the best census to build a model from. My preliminary theory is the model will need to be built over multiple stages with a combination of them, so that if this is the data used a first run with Ottoman household breakdowns from the 1840s and a second with the British census from 1890 would be the best bet.[36] Further supposition is impossible without knowledge of the specifics of the machine learning technique to be used.

Another potential proxy is census data from China. Although thousands of years of sporadic data exists, "These statistics are full of faults which make it obviously impossible to put much confidence in them as measures either of the exact size of the population at any time or of its changes during any period. Their ups and downs are often patently incredible, and the numbers of persons and households are sometimes inconsistent."[37] A debate as rich as the low estimate/high estimate for the population of the Roman Empire about who exactly is counted when exists as well, the specifics which are too much to get into here. Suffice it to say, potential proxies exist from Chinese data as well, but in selecting one carries with it all the caveats of the

[35] Cuno, Kenneth M. and Michael J. Reimer "*The Census Registers of Nineteenth-Century Egypt: A New Source for Social Historians*" (*British Journal of Middle Eastern Studies*, Vol. 24, No. 2, 1997), 193-216
[36] Cuno and Reimer (1997), 196
[37] Durand, John D. "*The Population Statistics of China, A.D. 2-1953*" (*Population Studies*, Vol. 13, No. 3, 1960), 210

Egyptian proxies and the additional problem of a completely different continent. The biggest advantage of the Chinese data is a population pre modern medicine and public health.

In short, there are not perfect proxies. These preliminary suppositions hint at the expertise that will be required to select the best proxy for the machine learning model. Machine learning is not capable of discerning the best proxy on its own, so the ancient historian must pick what he thinks is best, aided only by feeding the program different subsets of data and observing its outputs. This is a nuanced and important job, but hopefully the ability to make many different models from data quickly will allow the creation of a better and more accurate model. Nineteenth century Egypt or Ancient China are logical first steps for the creation of the model, but as with any ancient demography sourcing a good proxy is fundamental for accurate analysis. All this merely reinforces the fundamental role the human element plays in machine learning, and the importance of proper proxy selection.

Conclusion

In 1996, Walter Scheidel wrote "as far as quantitative analysis is concerned, the end of ancient demography is already in sight."[38] Five years later, he reiterated the statement, writing "I stand by my conclusion: 'in sight' is not yet 'here'."[39] Unfortunately Scheidel, in his haste to take the results of his findings and apply them to new studies, closed the book too quickly on ancient demography. The end of is not in sight, because Scheidel isn't thinking big enough. While traditional statistical modeling has done about much as it can do for the field, at least with

[38] Scheidel, Walter "*Measuring sex, age and death in the Roman empire: explorations in ancient demography*" (*Journal of Roman Archaeology*, Supplementary Series, No. 21, 1996), 169-184
[39] Scheidel (2001), xxvi

the data set we have, an ongoing revolution in data science, specifically in machine learning, will immeasurably help ancient demography in the near future. The rise of machine learning will broaden our understanding of the census returns, as it allows the building of high quality predictive models that far outstrip any previous methods. Equally important is the ease with which it will be able to be deployed. No longer will historians be forced to rely on the statistical analysis of a few in the field, but will rather be able to deploy machine learning to the data sets of their choice, broadening the discussion and allowing the testing of many different historical proxies at a far lower cost of time and energy.

Developing the R code necessary to unleash this revolution is beyond the scope of this paper, as no one has done it yet.[40] In the early stages of this hyper-developing field, scientists are even now honing applicable techniques. To completely unlock the secrets of the census returns from Roman Egypt will require intense collaboration between a skilled ancient historian and a skilled machine learner. The insight they will reach will be greater than anything they would have reached on their own, or anything possible from traditional statistical modeling. The burgeoning field of machine learning offers an exciting opportunity to revisit existing data with greater powers of prediction and superior proxies, and woe to the ancient demographer who ignores the opportunity.

[40] I looked for the R code from the first paper, the one that used census data from 2007-2012, and could not find it. The paper supplied all the raw data and a supplementary table, but no raw R code. While I am disappointed I am unable to write the R code myself, the purpose of the paper is really to show that this kind of analysis can be done in the future. I had not considered the idea of outsourcing the code writing internationally, that option might allow a skilled ancient historian with the right connections to write this paper themselves. I would still be wary however, as I think the combined challenges of both small data and demography are above and beyond a typical freelance job.