Comparing Artificial Neural Network and Cohort-Component Models for Population Forecasts

Viktoria Riiman*[1], Amalee Wilson†[2], Reed Milewicz‡[3], and Peter Pirkelbauer§[2]

[1]Center for Business and Economic Research, University of Alabama
[2]Computer and Information Sciences Department, University of Alabama at Birmingham
[3]Sandia National Laboratories

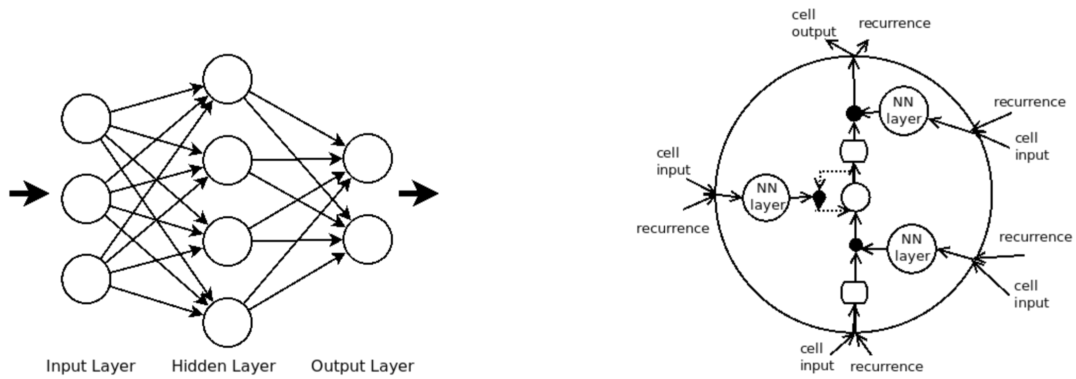
Abstract

Artificial neural network (ANN) models are rarely used to forecast population in spite of their growing prominence in other fields. We compare the forecasts generated by ANN long short-term memory models (LSTM) with population projections from traditional cohort-component method (CCM) for counties in Alabama. The evaluation includes forecasts for all 67 counties that offer diversity in terms of population and socioeconomic characteristics. When comparing projected values with total population counts from the 2010 decennial census, the CCM used by the Center for Business and Economic Research at the University of Alabama in 2001 produced more accurate results than a basic multi-county ANN LSTM model. Only when we use single-county models or proxy for a forecaster's experience and personal judgment with potential economic forecasts, results from ANN models improve. The results indicate the significance of forecaster's experience and judgment for CCM and difficulty, but not impossibility of substituting these insights with available data.

1 Introduction

Artificial neural networks (ANN) are frequently used for forecasting in numerous domains (Crone et al. (2011)), such as finance (Niaki and Hoseinzade (2013)), biology (Chon et al. (2000)), and tourism (Claveria et al. (2015)). However, not many attempts at using them for population estimates and projections have been made so far (Nordbotten (1996); Tang et al. (2006); Folorunso et al. (2010)). This is despite the fact that the potential for using ANN models in projecting population was noted more than a decade ago (Smith et al. (2001)). Prior use of ANN

*vriiman@cba.ua.edu
†amalee@uab.edu
‡rmilewi@sandia.gov
§pirkelbauer@uab.edu



(a) A neural network

(b) A cell in an LSTM network

Figure 1: Overview of Neural Network

models have utilized feed-forward networks with back propagation (Folorunso et al. (2010); Nordbotten (1996)) or fuzzy networks (Tang et al. (2006)). These models were shown to perform better than ratio correlation regression models for projecting population (Tang et al. (2006)) and better than forecasts that plug projected fertility, mortality, and migration data into a cohort-component equation (Folorunso et al. (2010)). Specifically, Tang et al. use data on birth, death, and school enrollment to compare their results with 2000 Census, while Folorunso et al. use 1990-2060 fertility, mortality, and migration data produced by the National Population Commission to compare their results with target population predictions.

In this paper, we forecast population using a long short-term memory (LSTM) network. LSTM models have become increasingly popular due to their ability to retain memory. We compare projection capabilities from our ANN LSTM models with the population projections developed at the Center for Business and Economic Research (CBER) at the University of Alabama in 2001. CBER forecaster, Carolyn Trent, projected population using the cohort-component method (CCM). We assess the accuracy of both methods by comparing them to actual population counts from the 2010 Census or mid-year population estimates by the U.S. Census Bureau.

To attain a high-quality cohort-component model, forecasters refine the forecasting methods using their experience. ANN models do not have that capability yet, hence we experiment by proxying the lack of cognitive ability with actual economic and demographic data. After experimenting with different types of models and training methods, the results showed that CCM, in general, provided better results than a basic multi-county ANN LSTM model. Using a single-county ANN LSTM model improved the results overall compared to CCM.

2 Artificial Neural Network Model - LSTM

Artificial neural networks (ANN) is a machine learning approach that attempts to simulate cognitive functions. Its characteristics allow for modeling nonlinear relationships between data points. As depicted by Fig. 1a, a simple ANN consists of layers of cells. Each cell receives input from cells of the preceding layer and sends its output to the cells of the following layer. A cell's transfer function computes a weighted combination of its input connection, and fires an output if it exceeds a certain threshold. Before we can utilize an ANN, we need to train it. Under supervised training, input data and expected output data are provided for the model. During the training phase, the ANN adjusts its cells' weights and thresholds to produce the desired output. The most popular technique for that is backpropagation (Folorunso et al. (2010)).

Several architectures of neural networks have been proposed (Lipton et al. (2015)). The simplest networks are feed-forward networks. Feed-forward networks have been applied to predicting time-series data (Tang and Fishwick (1993); Claveria et al. (2015)), including population forecasts (Folorunso et al. (2010)). A refinement of feed-forward networks, fuzzy networks have also been applied to population projection (Tang et al. (2006)).

A drawback of forward networks is the lack of memory, as cells do not have ability to remember previous outputs. A recurrent network model (RNN) provides the notion of memory. An RNN's cell output of time t is fed back as an input at the time step $t+1$ to the same cell. Thereby, an RNN can retain information for an infinite amount of time. This makes it a powerful model (Siegelmann and Sontag (1995)), useful to time-series prediction (Hüsken and Stagge (2003); Claveria et al. (2015)).

While RNNs introduce the notion of memory, training a network to retain information for a long time is difficult (Hochreiter (1991); Bengio et al. (1994); Hochreiter et al. (2001)). To overcome this challenge, long short-term memory network (LSTM) have been proposed (Hochreiter and Schmidhuber (1997); Gers et al. (2001)). LSTM enhances RNN models with the addition of long short-term memory. Thus, the cell can take its history into account. At any given time-step a cell can selectively choose to forget or replace some of the memory. Fig. 1b shows an example of one of the LSTM cell designs. An LSTM's input is comprised of the cell's input and the recurrence of its output at the previous time step. The content of LSTM's memory is controlled by internal layers that supervise how much old information is retained and what new information is added. As a result, the output is computed taking all this information into account.

3 Cohort-Component Method

The cohort-component method is a traditional forecasting approach in demography (Smith et al. (2013)). It is also the most popular method among the members of the U.S. Census Bureau Federal-State Cooperative for Population Projections (FSCPP). According to the 2015 FSCPP survey, 75% of FSCPP members use cohort-component method based on historical demographic data (Hunsinger (2015)). The next two popular methods were used by 27.5% and 22.5% of respondents, respectively: trend extrapolation of total population data and top-down methods such as constant-share, shift-share, and share of growth.

The Center for Business and Economic Research at the University of Alabama, established in 1930, has a long history of developing population projections using cohort-component method. The Center forecasters use bridged-race population estimates from the Centers for Disease Control and Prevention (CDC), specified in five-year age groups from 0-4 through 80-84 (Center for Business and Economic Research (CBER) (2001)). For computational purposes, the 0-4 age group is split into under 1 and 1-4 components, while individuals 85 and over are grouped in a single category. Breaking each age group down by race and gender yields 76 age/race/gender cohorts (using two race groups: white population, as well as black and others population).

The basic equation of the projections is:

$$P_t = P_{t-1} + B_{t-1,t} - D_{t-1,t} + M_{t-1,t}$$

where, P_t refers to population at time t; P_{t-1} population at time $t-1$; $B_{t-1,t}$ the number of births in the interval from time $t-1$ to time t; $D_{t-1,t}$ the deaths in the interval from time $t-1$ to time t; and $M_{t-1,t}$ is the net migration in the interval from time $t-1$ to time t, which equals in-migration minus out-migration.

The projection process is carried out in five-year increments, run independently for each geography. The calculation of birth, death, and migration components forms the basis for these projections. Additionally, multiple counties are adjusted for institutional effects such as colleges, universities, military installations, prisons, and nursing homes. Generally, assumption is made that populations in these institutions would not change in size or age distribution throughout the projected period.

Overall, discussions of forecasting being an art as well as a science are common among forecasters. Projecting population is noted to be an art that is influenced by scientific techniques and that personal opinions, judgments, experience, and outlook are used throughout the process (Guimarães (2014); Daponte et al. (1997)). Thus, CBER forecaster's experience and personal judgment in making assumptions for birth, death, and migration components are important for the accuracy of projections. Moreover, forecaster's opinions are also used when verifying that projected total population and age/race/gender distributions for counties make sense given available information about planned economic developments, potential formation of new school districts, expected changes to prison populations, possible army personnel movements, and all other useful local knowledge.

4 Empirical Evaluation

4.1 Data Description

For ANN LSTM models we used data available for all 67 Alabama counties: mid-year intercensal population estimates from 1969, developed by the U.S. Census Bureau. We also use decennial census data by county for each census year between 1910 and 2010, available from the National Historical Geographic Information System. The former is the first year when the data for all 67 Alabama counties are available.

For selected models we also use births and deaths data from the Center for Health Statistics at the Alabama Department of Public Health. Additionally, economic data are used from the Bureau of Economic Analysis such as proprietors' employment, wage and salary employment, real per capita income, and real average earnings per job. Births, deaths, and economic data as mid-year population are available from 1969. In addition, we use dummies for economic development from 2016 Alabama Workforce Development Councils that divide the state territory into 10 geographically compact regions.

Alabama counties offer diversity in terms of total population, population dynamics, and socioeconomic characteristics. Table 1 provides an overview of data range for total population in 1910 and 2010; for population growth rates between the decennial censuses; for births, deaths, and net migration (that is estimated by subtracting births and deaths impacts from change in total population) in 2010; and for real per capita income and average earnings in 2010.

Since CBER projections using cohort-component method for 2010 were based on data up to and including 2000, in order to have equivalent projections for comparison, we use only the data up to and including 2000 for most ANN LSTM models as well. In order to try to substitute for forecaster's knowledge on upcoming events affecting population such as economic and housing developments,

Table 1: Data Overview

	State of Alabama	County - min	County - max	County - median
Population 1910	2,138,093	12,855	226,476	27,155
Population 2010	4,779,736	9,045	658,466	34,339
Population growth:				
1910-20	9.8%	−20.3%	36.9%	7.8%
1920-30	12.7%	−21.5%	39.2%	4.9%
1930-40	7.1%	−5.0%	30.7%	5.2%
1940-50	8.1%	−23.2%	62.8%	−4.8%
1950-60	6.7%	−21.5%	61.0%	−7.1%
1960-70	5.4%	−21.7%	70.4%	0.6%
1970-80	13.1%	−10.4%	74.3%	10.6%
1980-90	3.8%	−15.0%	49.9%	−0.7%
1990-2000	10.1%	−8.5%	44.2%	7.6%
2000-10	7.5%	−16.1%	36.1%	0.9%
Births 2010	59,979	86	8,883	419
Deaths 2010	47,897	90	6,773	427
Net migration 2010-11	3,791	−2,485	3,105	−53
Real per capita income 2010	$33,510	$22,656	$42,164	$28,775
Real average earnings per job 2010	$43,472	$27,159	$61,343	$34,241

Source: U.S. Census Bureau, Center for Health Statistics, Alabama Department of Public Health, Bureau of Economic Analysis, and Center for Business and Economic Research, The University of Alabama.

in some models we add economic and/or births and deaths data for 2001-2010. Though it is not feasible to have such perfect projections, we added them to check the significance of having these data for forecasters.

4.2 ANN LSTM Model Specification

We based our model on a reference implementation for time-series prediction (Brownlee (2016)) and implemented it in Keras, a Python package that uses the open-source TensorFlow software library (Géron (2017)). All developed models consisted of an LSTM layer (§2) and a Dense layer from Keras. While the LSTM layer is responsible for most of the work associated with learning the time series prediction, the Dense layer is responsible for producing a single network output from the LSTM layer's output. While the Dense layer is a regular, densely connected network of neurons with a linear activation function, the LSTM layer consists of nonlinear functions. Tests were run using more complicated networks, with more and different layers, but it was found that relatively simple models perform well for our analysis.

In order to explore trade-offs of different approaches to project data we developed two different variations. (1) Model A, multi-county model: the question that we attempted to answer with this model is whether we can gain better insights into population dynamics by training a single model from data available for all counties. Under this scenario, input was normalized across all counties, which is necessary for maintaining the relative difference between small and large counties. Once the model was trained, on one county data at a time, it was repeatedly used to predict each single county's output for the projected years. (2) Model B, single-county model: this is a set of models, one per county, each of which was trained separately on that county data only. All trained models share the same specifications. Thus, each model was specialized for projecting a county's mid-year or decennial census population independently from other counties' data. Any input data and training data was normalized to values between 0 and 1 for a single county.

Both ANN LSTM models offer various parameters that allowed us to experiment with different setups. Since a neural network is repeatedly trained on the same data over n epochs, we experimented with different numbers of epochs up to 200 (using batch size of one). Further, we used different window sizes over a number of past years that are fed into the network to project the population at the next point in time. This allows to account for potential autoregressive processes in population data. Another parameter we experimented with was network size, from four to 32 LSTM cells. We also experimented with different loss functions. The mean absolute error function provided overall best results. We relied on the Keras default activation functions and used the adam optimizer (Kingma and Ba (2014)). We also experimented with different sizes of training and validation sets while keeping test sets for decennial census as 2010 and for mid-year population as 2001-2010 data points.

We experimented with two different input data, decennial census and mid-year annual population data, applying one-step-ahead forecasting approach. For the decennial census data, we trained the model for the period between 1910 and 2000, and then we projected 2010. For these data, we started with a time window of size one, essentially basing the projection on the last measured population ten years earlier. Increasing the window size produced worse results for most counties. The mid-year population was projected over a ten year period, one year at a time. Thus, the output/forecast at time step t_1 was appended to the input data for the next time step t_2. Under this scenario, it could be possible that early projection errors amplify over the ten year period. For the mid-year population model, we found that a prediction window of five gives good results (though for some specific counties the best results varied). Regarding LSTM cells and number of epochs used for training, the best results for Model A with mid-year population had 5 LSTM cells and 10 epochs and with decennial census had 4 LSTM cells and 100 epochs, while for Model B with both types of population data 16 LSTM cells and 100 epochs provided best results.

4.3 Results

Since previous papers comparing ANN with other methods used for population projections showed the results favorable for ANN, we expected a similar outcome when comparing those with projections from utilizing cohort-component method in Alabama. However, when running a basic multi-county model, ANN LSTM model A, we found that the cohort-component method yields more accurate results displayed in Table 2, except for mean absolute percent error (MAPE) when decennial census data were used. Thus, CCM and forecaster's experience and personal judgment were indeed important for the accuracy of results.

Only when we used a single county model, ANN LSTM Model B, the results improved. Using decennial census data produced better results for all three errors: root mean-squared error (RMSE), mean absolute error (MAE), and MAPE. Using mid-year population data produced smaller MAPE, but still had larger RMSE and MAE.

We explored substituting forecaster's experience and personal judgment with true data for births, deaths, and economic data during 2001-2010 period (Table 3). Although it is not feasible to have such accurate data, the experiments offer insight into the importance of these data for projec-

Table 2: Comparison of Estimate Errors: Cohort-Component vs ANN LSTM, 2010 Data

Method	**RMSE**	**MAE**	**MAPE**
Cohort-Component Method	5,251	3,216	6.5%
ANN-LSTM Model A			
Mid-year population	17,523	11,004	16.7%
Decennial census population	5,442	3,346	6.3%
ANN-LSTM Model B			
Mid-year population	7,160	3,663	6.3%
Decennial census population	4,529	2,742	5.0%

Model A: Model is trained on data from all counties, with training process done on one county data at a time.
Model B: Each county has a separate model trained on data from that county only. All 67 counties have the same model specification.
RMSE: root mean-squared error, *MAE*: mean absolute error, *MAPE*: mean absolute percent error

tions. Because the data were available since 1969, we used it for projecting mid-year population. The results showed that true economic data improved the results in MAPE from 6.3% to 5.0%, but true births and deaths data increased it to 8.8%, though RMSE has decreased. Adding births and deaths to economic data improved RMSE and MAE, but slightly increased MAPE.

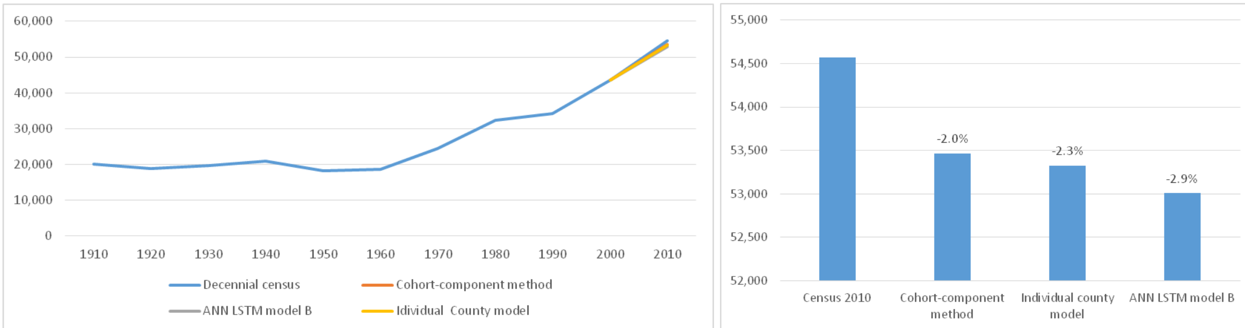
Table 3: Estimate Errors Using True Data in ANN LSTM Model B, 2010 Comparison

Model	**RMSE**	**MAE**	**MAPE**
True births and deaths	6,273	4,215	8.8%
True economic data	5,984	3,196	5.0%
True births, deaths, and economic data	4,844	2,988	5.1%

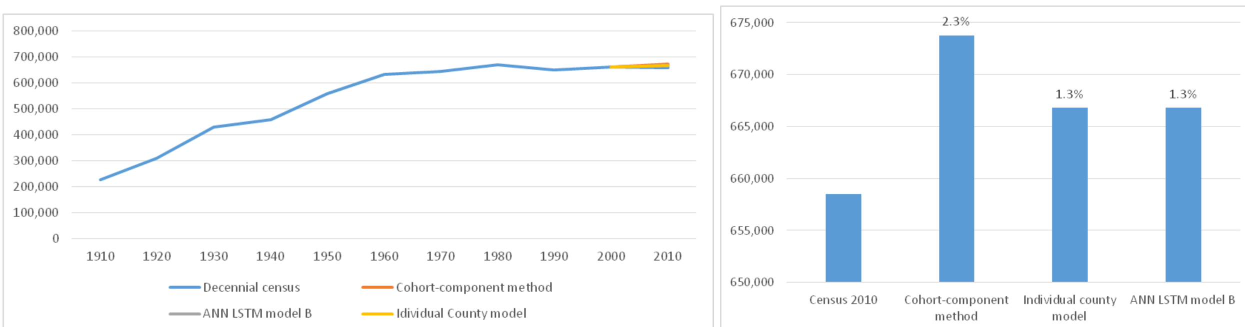
Note: Comparison of projected mid-year 2010 population and 2010 mid-year population estimate from the U.S. Census Bureau.

Since ANN LSTM model B only allows the same specifications for all the counties, we experimented how optimizing a model for an individual county could yield better results. Indeed, choosing specifications to fit a particular county sometimes showed better results than ANN LSTM model B. For Autauga County, for example, a model with individual specification showed smaller errors, with projected population being 2.3% below the true value compared to 2.9% (Fig. 2). The CCM model still produced better results with projections being 2.0% lower than the actual 2010 Census. On the other hand, for Jefferson County, the most populous county in the state, ANN LSTM Model B provided the best results as we did not find any optimization for that county. LSTM model projections were better than the CCM projections for Jefferson County.

We also examined how models handled the most populous and the least populous counties. When looking at the best performing model, ANN LSTM Model B that uses decennial census population data, eight out of ten most populous counties were among the top ten counties with largest RMSE and MAE, none were among the top ten counties with largest MAPE, and two counties were among the bottom ten counties with smallest MAPE (Table 4). In comparison, when



(a) Autauga County: Population and 2010 Projections



(b) Jefferson County: Population and 2010 Projections

Figure 2: Population and 2010 Projections

Note: Percentages indicate differences between projections and 2010 Census true values.

ANN LSTM Model B with intercensal mid-year population data, the worst performing model, was used, five out of top ten most populous counties were among the top ten counties with largest RMSE and MAE, one county was among the top ten counties with largest MAPE, and two counties were among the bottom ten counties with smallest MAPE.

5 Discussion

Forecaster's experience and personal judgment seem to have a strong impact on the accuracy of population projections since results from the cohort-component method were, in general, better than from the ANN LSTM model A, which was trained on data from all counties, one county data at a time. Results from ANN LSTM model B, which used single county data when training the model, gave better results than the CCM when decennial census data were used. When mid-year population data were used for model B, it gave better MAPE but worse RMSE and MAE than the CCM. Thus, we may still need to find ways to substitute for forecasters' personal experience, judgment, and information available to them when developing population projections. Using more data, such as economic forecasts, could be one such option.

Training ANN LSTM model only on the data from the county for which projections are later made gave better results, indicating that population development trends in other counties did not affect a specific county's population projection as much as a county's own historical trends and instead may have created additional noise for the ANN LSTM model. The diversity of population

Table 4: Error Rankings by County from the Best and Worst Performing ANN LSTM Models

		ANN LSTM Model B, decennial census population		ANN LSTM Model A, mid-year population	
County	2010 Census population, rank	RMSE & MAE, rank	MAPE, rank	RMSE & MAE, rank	MAPE, rank
Jefferson	1	5	59	11	63
Mobile	2	41	67	2	36
Madison	3	1	27	16	59
Montgomery	4	9	44	17	52
Shelby	5	7	36	1	8
Tuscaloosa	6	2	11	5	31
Baldwin	7	4	30	3	11
Lee	8	3	12	4	12
Morgan	9	17	46	12	29
Calhoun	10	8	24	18	41
Crenshaw	58	56	41	54	43
Choctaw	59	29	4	43	18
Sumter	60	51	32	47	23
Conecuh	61	43	21	48	21
Wilcox	62	52	28	40	7
Coosa	63	67	58	49	19
Lowndes	64	32	2	35	3
Bullock	65	66	52	42	10
Perry	66	54	29	37	2
Greene	67	46	13	34	1

Note: The results are from the best performing (Model B, decennial census data) and worst performing model (Model A, mid-year intercensal data) for ten most populous and ten least populous counties. Errors are ranked from the largest (rank #1) to the smallest error (rank #67).

dynamics among Alabama counties may have been the reason for this noise. Adding more geospatial data to this model could improve the results, capturing the potential impact of geographic location on population growth. Although our experiments with adding economic development region dummies to the model A did not improve the results, further exploration is needed.

Using decennial census data for ANN LSTM models resulted in smaller errors than in models with mid-year population. Thus, having a longer time span as input produced better results. This could be caused by the length of projections that required ten steps for mid-year population compared with one step for decennial census data due to the nature of one-step-ahead forecasts. Using other forecasting techniques, such as direct prediction (Bontempi (2008)), could improve projections of mid-year population.

In this work, we only experimented with a small set of parameters of a reference implementation. It is our goal to explore automated techniques for finding better optimized ANN LSTM models (Goodfellow et al. (2016)). With the continuing development of ANN models, we are expecting to receive improved forecasts.

Overall, using ANN models to project only some population components instead of total population could be worthwhile to explore in the future. This could make ANN models another alternative tool in the toolbox of demographers.

Acknowledgement

We are grateful for the constructive comments from Ron Prevost.
This work was run on computer equipment purchased using NSF grants CNS-0821497 and CNS-1229282.

References

Bengio, Y., Simard, P., Frasconi, P., 1994. Learning long-term dependencies with gradient descent is difficult. Trans. Neur. Netw. 5(2), 157–166.

Bontempi, G., 2008. Long term time series prediction with multi-input multi-output local learning. In: 2nd European Symposium on Time Series Prediction (TSP), ESTSP'08, pp. pp. 145–154.

Brownlee, J., 2016. Time series prediction with LSTM recurrent neural networks in python with keras. `http://machinelearningmastery.com/time-series-prediction-lstm-recurrent-neural-networks-python-keras/`, accessed on December 19, 2016.

Center for Business and Economic Research (CBER), 2001. Alabama Population Projections 2010-2040. Methodology. The University of Alabama, Culverhouse College of Commerce.

Chon, T.-S., Park, Y.-S., Kim, J.-M., Lee, B.-Y., Chung, Y.-J., Kim, Y., 2000. Use of an artificial neural network to predict population dynamics of the forestpest pine needle gall midge (diptera: Cecidomyiida). Environmental Entomology 29(6), 1208.

Claveria, O., Monte, E., Torra, S., 2015. Multiple-input multiple-output vs. single-input single-output neural network forecasting. Working paper 2015/02 1/28, Research Institute of Applied Economics, Regional Quantitative Analysis Research Group.

Crone, S. F., Hibon, M., Nikolopoulos, K., 2011. Advances in forecasting with neural networks? empirical evidence from the {NN3} competition on time series prediction. International Journal of Forecasting 27(3), 635 – 660, special Section 1: Forecasting with Artificial Neural Networks and Computational IntelligenceSpecial Section 2: Tourism Forecasting.

Daponte, B. O., Kadane, J. B., Wolfson, L. J., 1997. Bayesian demography: projecting the iraqi kurdish population, 1977-1990. Journal of the American Statistical Association 92(440), 1256–1267.

Folorunso, O., Akinwale, A., Asiribo, O., Adeyemo, T., 2010. Population prediction using artificial neural network. African Journal of Mathematics and Computer Science Research 3(8), 155–162.

Géron, A., 2017. Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. O'Reilly Media, 1st edition.

Gers, F., Eck, D., Schmidhuber, J., 2001. Applying lstm to time series predictable through time-window approaches. In: Proceedings of the International Conference on Artificial Neural Networks, ICANN '01, pp. 669–676, Springer-Verlag, London, UK, UK.

Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. MIT Press, `http://www.deeplearningbook.org`.

Guimarães, R. R. d. M., 2014. Uncertainty in population projections: the state of the art. Revista Brasileira de Estudos de Populaçõ 31, 277–290.

Hochreiter, S., 1991. Untersuchungen zu dynamischen neuronalen Netzen. Diploma thesis, Institut für Informatik, Lehrstuhl Prof. Brauer, Technische Universität München.

Hochreiter, S., Bengio, Y., Frasconi, P., Schmidhuber, J., 2001. Gradient flow in recurrent nets: the difficulty of learning long-term dependencies. In: Kolen, J., Kremer, S. (Eds.), Field Guide to Dynamical Recurrent Networks, IEEE Press.

Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9(8), 1735–1780.

Hunsinger, E., 2015. FSCPP Members 2015 Questionnaire. Alaska Department of Labor and Workforce Development, State Demographer's Office.

Hüsken, M., Stagge, P., 2003. Recurrent neural networks for time series classification. Neurocomputing 50, 223–235.

Kingma, D. P., Ba, J., 2014. Adam: A method for stochastic optimization. CoRR abs/1412.6980.

Lipton, Z. C., Berkowitz, J., Elkan, C., 2015. A critical review of recurrent neural networks for sequence learning. CoRR abs/1506.00019.

Niaki, S. T. A., Hoseinzade, S., 2013. Forecasting S&P 500 index using artificial neural networks and design of experiments. Journal of Industrial Engineering International 9(1), 1.

Nordbotten, S., 1996. Neural network imputation applied to the Norwegian 1990 population census data. Journal of Official Statistics 12(4), 385–401.

Siegelmann, H., Sontag, E., 1995. On the computational power of neural nets. Journal of Computer and System Sciences 50(1), 132–150.

Smith, S. K., Tayman, J., Swanson, D., 2001. State and Local Population Projections. Methodology and Analysis. Kluwer Academic/Plenum Publishers, NY.

Smith, S. K., Tayman, J., Swanson, D., 2013. A Practitioner's Guide to State and Local Population Projections. Springer Media Dordrecht.

Tang, Z., Fishwick, P. A., 1993. Feed-forward neural nets as models for time series forecasting. ORSA Journal of Computing 5(4), 374–386.

Tang, Z., Leung, C. W., Bagchi, K., 2006. Improving population estimation with neural network models. In: Proceedings of the Third International Conference on Advances in Neural Networks - Volume Part III, ISNN'06, pp. 1181–1186, Springer-Verlag, Berlin, Heidelberg.