

An Analog Method for Seasonal Forecasting in Northern High Latitudes

John E. Walsh^{1*}, Brian Brettschneider², Nathan P. Kettle¹, Richard L. Thoman¹

¹International Arctic Research Center, University of Alaska Fairbanks, Fairbanks, Alaska, USA ²National Weather Service Alaska Region, NOAA, Anchorage, Alaska, USA Email: *jewalsh@alaska.edu

How to cite this paper: Walsh, J.E., Brettschneider, B., Kettle, N.P. and Thoman, R.L. (2021) An Analog Method for Seasonal Forecasting in Northern High Latitudes. *Atmospheric and Climate Sciences*, **11**, 469-485. https://doi.org/10.4236/acs.2021.113028

Received: April 30, 2021 **Accepted:** June 12, 2021 **Published:** June 15, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0). http://creativecommons.org/licenses/by/4.0/

cc ① Open Access

Abstract

An analog forecast method designed for monthly and seasonal outlooks is applied to the Arctic. The analog selection process uses pattern matches based on agreement with historical data to identify past years with similar distributions of sea level pressure, upper-air geopotential height, surface and upper-air temperatures, precipitation, and sea surface temperatures. The evolution of the atmosphere in the analog years is then the basis of a prediction for the target year. Users can choose the predictor domain, the predictand domain, the variable to be predicted, and the number of antecedent months on which the analog selection is based. We provide an example of a monthly forecast generated by the analog forecast tool. In comparisons with operational dynamical model forecasts over the period 2012-2019, the analog system underperforms the dynamical models in middle latitudes but generally outperforms the dynamical models in monthly forecasts of surface air temperatures in the Arctic. The improvement over the dynamical models is especially apparent in the late summer and early autumn (August-October).

Keywords

Arctic, Climate, Analog, Forecasting

1. Introduction

While the concept of analog forecasting has been known since the advent of weather forecasting, serious scholarly research on the use of analogs for weather and climate forecasting began in the late 1960s and into the 1980s [1] [2] [3] [4] [5]. These studies, and many others, sought to identify large-scale patterns in the atmosphere and ocean as a guide to forecasting weather in the over time-scales ranging from days to seasons. On the seasonal time scale, perhaps the

most well-known analog forecasts are derived from El Niño and La Niña states in the tropical Pacific Ocean. Maps showing what *typically* happens during an El Niño and a La Niña proliferate on the Internet [6]. An example is shown in **Figure 1**, in which panel (a) corresponds to El Niño conditions and panel (b) corresponds to La Niña conditions. Such patterns enable seasonal climate outlooks via a two-step process. Step 1 is the identification of an El Niño or La Niña event in the tropical Pacific that is expected to persist through the next season (or longer). Step 2 is the use of the broad patterns of association in **Figure 1** as the basis for a seasonal forecast for regions outside the tropics. In essence, the previous El Niño or La Niña years are being used as analogs for a forecast from current conditions.

Current analog forecasting initiatives fall into one of several categories. In the first, pattern marching or similar techniques are used to identify similar past cases that are then followed into the future as a forecast [7] [8] [9]. These methods



Climate Prediction Center/NCEP/NWS

Figure 1. Winter anomalies of temperature (red = warmer than normal, blue = cooler than normal) and precipitation (green = wetter than normal, brown = drier than normal) in winters with (a) El Nino and (b) La Nina conditions in the equatorial Pacific. Characteristic jet stream patterns are shown by thick arrows. From Climate Prediction Center/NCEP/NWS/NOAA, https://www.cpc.ncep.noaa.gov/products/analysis monitoring/ensocycle/enso_cycle.shtml.

usually assume no a priori knowledge of the climate system. A second approach is to utilize analogs in conjunction with dynamical (numerical weather prediction, NWP) models [9] [10]. More recently, machine learning techniques have been brought to bear on the use of analogs in weather forecasting, particularly in the targeting of surface temperature extremes over multiday time scales [11] [12]. Despite this upswing of research interest, analog-based methods have not been prominent in weather and climate forecasting strategies in the past few decades, especially with the advances in dynamical models for weather forecasting and climate simulations. For purposes of deterministic weather forecasts, the conventional wisdom is that analogs diverge sufficiently rapidly that a forecast based on analogs cannot compete with forecasts from dynamical models. The divergence arises from nonlinear error growth triggered by differences between any two analog states. However, dynamical models are also sensitive to errors in the initial conditions, and several additional considerations suggest that a fresh look at an analog approach may be in order, especially at monthly-to-seasonal lead times. These considerations include the recognition that slowly evolving surface boundary conditions (sea surface temperature, sea ice, soil moisture) contribute to departures from normal atmospheric states. Potential applications are especially deserving of exploration for regions such as the Arctic, where the analog method has not previously been evaluated, where observational data to initialize dynamical models are relatively sparse, and where dynamical model skill is lower than in middle latitudes

(https://www.ecmwf.int/en/about/media-centre/science-blog/2018/improving-pr ediction-and-climate-monitoring-polar-regions).

Additional motivation for the present study includes the following. First, beyond the range of several weeks, the skill of dynamical models in predicting departures from monthly and seasonal means is small. Second, statistical approaches such as those used by the Climate Prediction Center for its seasonal outlooks (e.g., CCA—Canonical Correlation Analysis, Regression Analysis (SMT) tool), implicitly assume that there is some longer-term forcing that can provide forecast skill. This forcing may arise from persistence in ocean temperatures or other parts of the land-ocean system, or from evolution of this forcing inherent in the analog approach. In this regard it should be noted that an analog approach (constructed analogs, CA) is part of the Climate Prediction Center's toolkit for seasonal outlooks

(http://www.cpc.ncep.noaa.gov/products/predictions/90day/seasglossary.html#c a). Third, analog approaches have recently been developed and made available for various parts of the contiguous United States, for example by the Cooperative Institute for Precipitation Systems (CIPS) at St. Louis University

(<u>http://www.eas.slu.edu/CIPS/ANALOG/analog.php</u>). While the CIPS analog forecast system is applied only to shorter forecast ranges (days), it 1) demonstrates the use of analog forecast targeting impacts, which in the CIPS case are tied to severe weather threats and 2) is publicly available.

The targeting of applications and the real-time availability highlight one of the main advantages of analog forecasts: stakeholders and decision-makers can be given historical analogs for use in assessing impacts of greatest relevance to their particular decision-making needs (e.g., related coastal flooding or, more generally, the likelihood of severe weather) by knowing which past events most closely resemble their present situation. On the seasonal timescale, analog years provide potentially useful information for anticipating wildfire season severity, droughts or, in the Arctic, sea ice conditions affecting navigation and other offshore activities. The fact that the analog forecast approach has not been rigorously explored for the Arctic represents an opportunity for enhancing monthly-to-seasonal outlooks in a region that is arguably under-served by the suite of existing forecast products.

The considerations summarized above led to the present study, for which the objectives were to: 1) develop an analog system to forecast atmospheric conditions in the Arctic over timescales of a month to several seasons, and 2) assess the skill of the analog forecasts relative to dynamical model forecasts. Because our analog system will target forecasts on the monthly to seasonal timescale, monthly fields will form the basis of the analog selection. The study is carried out with the aid of software that gives the user the choice of region, forecast lead time, and predictor variables, all at monthly temporal resolution.

2. Data and Methods

2.1. Data Source

The analog forecast procedure presented here uses current atmospheric and sea surface temperature to from the NCEP/NCAR R1 reanalysis [13] [14] to identify a set of n best historical matches (analogs). For illustrative purposes, we choose n = 5. A previous hindcast study [15] has shown that values of n in the range of 5 to 10 represent an optimal compromise between excessive damping of anomalies as the number of analogs increases above about 5 - 10 and the increased volatility as the number of analogs is reduced below five. The R1 reanalysis is used because it offers a longer period of record (since 1949) than most other atmospheric reanalyses and is routinely updated (to within a day or two of real time). Its resolution of $2.5^{\circ} \times 2.5^{\circ}$ in latitude and longitude enables the running of the required matching algorithm in less than a minute on a laptop or desktop computer.

2.2. Generation of the Analog Forecast

The analog identification is based on six variables from an antecedent period: surface air temperature, sea level pressure, precipitation, upper–air pressure (geopotential height), upper–air temperature, and sea surface temperature. As described below, there are several options for the pressure level of the upper-air geopotential height and temperature. The antecedent period, which is also user-selected, can be as current as the most recent calendar month. The main attributes of the analog system are presented in **Figure 2**. In order to capture the



Figure 2. Schematic representation of the analog forecast system.

essence of the procedure by which a user obtains an analog-based forecast, we highlight the main elements of **Figure 2** by stepping through a user's construction of a forecast.

First, a search area and calendar month(s) for which the analog years are selected by the user (upper left panel in **Figure 2**). Monthly data are available, so the time-window can be specified as a single month (e.g., January, 2020) or a set of 2, 3, ..., 12 consecutive months (e.g., June through August 2020). The geographical area should be one that the user believes is climatologically predictive of the intended forecast area. One natural choice is a region in the Pacific Ocean near the equator which is part of teleconnection patterns (El Niño, La Niña) known to be associated with monthly-to-seasonal variations over a large area (**Figure 1**). The user also selects the pressure levels (925, 500 or 250 hPa) for the upper air geopotential height and the air temperature fields to be used in the analog-year selection. The upper-air fields used in the analog selection are in addition to the near-surface fields of sea level pressure, 2-meter air temperature, precipitation and sea surface temperature.

There are varying degrees of relationship among the user options. First, the geographical area of the predictors can be the same as the area for which the prediction is desired, although it need not be. In the left portion of **Figure 3**, the two areas are different. With regard to the variables, the six predictor variables are pre-determined or "hard-wired" into the tool, while the variable being pre-dicted is left to the choice of the user. The predicted variable can be any one of the six variables in the predictor suite. Both the predictor (analog selection) time-

frame and the prediction timeframe can range in length from 1 to 12 months. However, the month(s) used the analog selection (and the predictor variables) are different from the month(s) of the prediction. The former must precede the latter, and there must be no overlap; otherwise the predictor information would extend into the timeframe for which the prediction is being made, in which case there would not be a true "prediction".

Second, a forecast area and time window must be defined. Similar to the search area, data are available for monthly and multi-month periods. The area for which the forecast is made should be user-relevant.

Third, given the choices summarized above, the software evaluates the pattern matches of each variable with corresponding fields of all past years over the user-selected domain of the predictors. The pattern-matching uses the RMSE metric, and a weighting procedure (Section 2.3) is applied to obtain an overall match score for each year in the database. The weights are the standardized metrics of closeness of fit. The top five years by match score are selected as the basis for the analog forecast (**Figure 2**, upper middle panel).

Spatial distributions of the forecast variable in the forecast area during the forecast month are then returned for each of the five best analog years based on the match scores calculated during Step 3. These conditions are presented in map form (upper right panel of **Figure 2**). The conditions returned during the top five match years are also averaged to produce a composite forecast for the forecast period (lower right panel of **Figure 2**). The composite forecast may be regarded as a "constructed analog" forecast [16] [17]. This is a prediction of the user-requested field in the forecast area during the forecast period based on conditions during the five best analog years in the search area.

2.3. Pattern Match Criteria

What constitutes an analog match is a fundamental question in this type of study. While an analog pattern match need to be perfect to be useful, the closeness of the match should align with the utility of an analog. The closeness of a match may be measured qualitatively, or quantitatively by using various statistical metrics. [1] describes a simple procedure using squared differences between atmospheric states and also a root mean square error (RMSE) metric. Other techniques involve pattern correlation, linear regression [16] or more sophisticated statistical algorithms [18].

Analog matches in the current study use a similar match technique as [1] [19] [20] and others namely, a ranked RMSE score. The target month/year is evaluated against the same calendar month in all other years on a grid-cell by grid-cell basis. The difference between the target month/year and the comparison month/years is squared to remove the sign. Prior to the cell-by-cell evaluation, we remove the target domain mean from the target month/year and separately remove each comparison month/year domain mean from each comparison month/year. Without this step, the highest RMSE matches often reflect the

overall deviation from climatology—not the best pattern matches. Metrics based on simple departures from the mean are not appropriate for the atmospheric circulation, which is determined by spatial gradients rather than by departures from the mean.

Root mean squared error (RMSE) between each past year and the search period is performed for each grid cell, then summed and weighted across each variable using an auto-weighting calculation based on a standardized closeness of the match of each variable. The sum of the weighted mean squared errors across all variables is the "Match Score". Low numerical values of the match score, i.e., small values of the summed RMSEs, correspond to high levels of similarity between that year and the search year.

2.4. The User Interface

Figure 3 is a schematic representation of a user interface illustrating the options available for obtaining an analog-based forecast for a particular region. **Figure 3** aligns with **Figure 2** in the sense that **Figure 3**'s left and middle columns correspond with the upper and lower left panels, respectively, of **Figure 2**. An example of the software's response to the "Run analog forecast" icon in **Figure 3** is presented in Section 3.1.

3. Results and Discussion

3.1. Sample Forecast

In order to illustrate the performance of the analog forecast software, we present the following example. In this case, the domains for the analog selection and the actual forecast are identical, 60° N - 90° N. The variable being forecast is the 2-meter air temperature. The month being forecast is April 2020, while March 2020 is the period used to identify the best analog match years. This period was chosen because April 2020 was characterized by extreme temperature anomalies, particularly the anomalous warmth that was observed over northern Eurasia. **Figure 4** shows

Analog match search area & time

The analog match search area is the spatial region that is analyzed for statistical matches. This defaults to a region in the South Pacific which was empirically determined to correlate well with Alaska. Longitudes go from 0-360E, and latitudes go from 0-90N.



Forecast theme, area, and time span

Forecast area defaults to approximately the spatial extent of Alaska. Longitudes go from 0-360E, and latitudes go from 0-90N.

Forecast theme



Run analog forecast

Clicking the button below will open a new window that will run the analog forecast.

▲ It may take up to three minutes for the results to be available. Leave the window open until the processes completes.

Run analog forecast

Figure 3. Analog software user options.



Observed surface air temperature anomalies



that temperatures during March 2020 were generally 1°C - 4°C warmer than normal over most of northern Eurasia and the seas to the north. During April, however, temperature anomalies of +5°C to +8°C developed over central Siberia, with anomalies exceeding 5°C extending across much of the Arctic Ocean, including the Chukchi Sea. This highly unusual springtime warmth contributed to an early sea ice retreat north of Siberia, leading to record-low sea ice in in the Kara and Laptev Seas during much of the summer [21]. The anomalous heat persisted into June, leading to large areas of wildfire and degraded air quality over much of northern Asia [21]. The change from March to April is sufficiently large that a simple forecast of anomaly persistence from March to April would have severely underestimated the magnitude of the warmth over Siberia.

Figure 5 shows the temperature anomaly fields during March of the five years identified as the best analogs for March of 2020. These years, determined by application of the analog selection software to the NCEP reanalysis data for 1948-2019, were (in order of decreasing match) 2002, 2015, 2011, 1975 and 2014. The preponderance of recent years in the set of five best matches is consistent with the recent Arctic warming and the fact that the heaviest weights were assigned by this search to sea surface temperature and the 2 m air temperature. While the different sample members have varying locations and magnitudes of the high-latitude warm anomalies, all show positive departures from normal somewhere in the vicinity of northern Asia.

Figure 6 shows the temperature anomaly fields for April of the five best analog years. The anomalies are defined here as the departures from the climatological



Five best analogs for March 2020

Figure 5. Departures from normal (1981-2010) air temperatures (°C) during March of the five analog years with the closest match to 2020. Closeness of the match decreases rom upper left to lower right.

mean values for 1981-2010 (the current "climatological normal" period), although we note that the candidate years for the analog selection range from 1949 to 2019. As was the case for March (**Figure 5**), across-sample variability is apparent in the April temperature fields for the same analog years. In this case the 2nd, 3rd and 5th best analogs (2015, 2011 and 2014) have positive anomalies exceeding 5°C over parts of northern Asia, while the other two analogs have weaker positive anomalies over northern Eurasia. The across-sample variability in **Figure 5** and **Figure 6** is typical of the analog selection based on the approximately 70-year sample used here. As noted in Section 2.1, the skill of the analog forecasts is generally highest when the number of analogs used for the composite ("constructed analog") prediction is in the range of 5 to 10 [15].

The effect of compositing is shown in **Figure 7**, where the mean of the five analog forecasts for April 2020 is compared with the actual temperature anomaly map. The spatial pattern in the Arctic is generally captured well, as positive anomalies extend from Siberia over much of the Arctic Ocean in both the analog forecast and the reanalysis map. The negative anomalies over much of northern, western and central Canada are also well forecast by the analogs. The major area



Forecasts for April 2020 based on five best analogs for March 2020

Figure 6. Departures from normal (1981-2010) air temperatures (°C) during April of the five years with the best March analog matches for 2020.



Surface air temperature anomalies: April 2020

Figure 7. Departures from normal temperature in April 2020 based on the NCEP/NCAR reanalysis (left) and the composite forecast of the five best analogs (right).

180

150W

of discrepancy is from northeastern Canada to Greenland, where the negative anomalies forecast by the analogs did not verify. Over Baffin Bay, for example, the observationally-based reanalysis shows positive departures from normal where the analog system had forecast large negative anomalies.

Apr 2020

2014

While the preceding example illustrates the analog system's ability to capture some, but not all, features of month-to-month evolution of Arctic temperatures, it represents a single case. Moreover, the analog forecast was not compared to other forecasts. In the following section, we draw upon a larger sample of forecasts and include comparisons with operational models used for monthly to seasonal climate outlooks.

3.2. Comparison with Dynamical Forecast Models

The analog tool produces forecasts for periods of a month to several seasons. These forecast ranges are the same as those spanned by climate prediction centers' dynamical forecast models, which are generally coupled atmosphere-oceanice-land models. The atmospheric components of these models are essentially the same models used in numerical weather prediction, for which there is demonstrable skill in day-to-day weather predictions out to one or two weeks. When these models are run beyond a week or two in order to generate monthly and seasonal outlooks, the forecasts target monthly or seasonally-averaged departures from climatological means-as do the analog forecasts in this study. The availability of dynamical model forecasts for recent decades enables comparisons of the skill of the analog and dynamical model forecasts. For the comparison here, we use the North American Multimodel Ensemble (NMME), a set of models used to generate forecasts to lead times of seven to eleven months. The NMME ensemble of models has evolved over time (cf. Table 1 in [22]). The current forecasts of the six primary ensemble members of the current NMME are accessible at https://www.cpc.ncep.noaa.gov/products/NMME/ One of the NMME models is the National Centers for Environmental Prediction's CFSv2 [23], which we include in single-model comparisons described below.

At the global scale, the dynamical model forecasts of temperatures and heights were found to outperform the analogs method for most forecast months at all time lags. For example, over the 2012-2017 period, the analogs' 2-meter forecasts of temperature for the Contiguous U.S. using matches at one-month lead time beat the NMME only 40% of the time, with the analogs' with the percentage ranging from 49% in July-September to 25% in October-December (Figure 8) (The seasons in Figure 8 correspond to those of the annual cycle of temperature over 70°N - 90°N, where the months of minimum and maximum temperature are February and August, respectively). This under-performance of the analogs relative to the dynamical models diminishes as one moves into high latitudes and, for the Arctic, the skill of the analog forecasts slightly exceeds that of the dynamical models. Figure 9 shows the percentages of cases in which the analog and dynamical model forecasts had the lower RMSE of Arctic temperature forecasts. Each panel in Figure 9 shows the results for forecasts targeting a particular calendar month (January, February, ..., December) over the period 2012-2017. Each pair of bars summarizes results for cases in which the number (m) of months used for the analog selection ranged from 1 to 12. Each bar also includes







Figure 9. Percentage of cases in which the lower RMSE of the monthly forecast of Arctic air temperature was achieved by the analog system (blue bars) and the NMME (orange bars). Each panel summarizes forecasts for a particular calendar month of the 2021-2017 period (January to December from upper left to lower right). Successive pairs of bars in each panel are for cases in which the analog selection was based on m = 1, 2, ..., 12 antecedent months.

a range of lead times from one month to six or seven months. Because of occasional corruptions of the archived NMME files, the sample sizes for each bar range from 34 to 37 cases (generally six lead times for six different years).

Figure 9 shows that each type of forecast (analog, NMME) outperformed the other in some calendar months. For a given calendar month, the relative performance of the analog system shows some sensitivity to the number of antecedent months, m, used in the analog selection. However, the choice of m does not

have a major impact on which of the two methods shows has the smaller error in a particular calendar month. While there is substantial month-to-month scatter in the relative performance of the two methods, an aggregation of the results by season (**Figure 8**) leads to the conclusion that the analog method slightly outperforms the NMME in this particular application (forecasts of 2-meter temperature over the domain 70°N - 90°N, which is essentially the Arctic Ocean. The 60°N - 90°N was used in the analog selection for these forecasts. The aggregations in **Figure 8** include all values of m for each of the three months in the season.

Figure 8 and **Figure 9** show that, except for the autumn season, the analogs outperform the NMME in forecasts of 2 m air temperature in the Arctic. This suggests that the analogs are capturing some features of the Arctic system that have sufficient memory to impact subsequent air temperatures. The fact that the NMME outperforms the analogs in the autumn is consistent with the large autumn sea ice anomalies of 2012-2017, which are incorporated into the dynamical models' initializations with subsequent large impacts on the surface air temperatures during the freeze-up period. In this respect, the dynamical models' forecasts of autumn air temperatures over the Arctic during these years should be "hard to beat".

As a further illustration of the relative performance of the analogs and the dynamical models, Figure 10 shows the lowest RMSE errors of the forecasts made by the NMME, CFSv2 and analog forecasts for each month of the 2012-2019 period. All forecasts in this case are one-month forecasts, and the number of antecedent months used to identify the best analogs is m = 1. The errors show a strong seasonal cycle, with the largest errors in winter and the smallest errors in summer, consistent with the seasonal cycle of variance of Arctic air temperatures (Przybylak, 2000). The color-coded symbols in Figure 10 indicate which of the three forecasts was the "winner" in the sense that it had the lowest RMSE: analog (blue), NMME (orange) and CFSv2 (green). In over half (50/87, or 57%) of the cases, the analog forecast had the smallest error. In the remaining 37 cases, the lowest errors were about evenly split among the NMME (19 cases) and the CFSv2 (18 cases). The latter two forecasts are not independent, as the CFSv2 is one of the NMME models. Consistent with Table 1, the analog forecasts in Fig**ure 9** generally fare most poorly during the autumn season (the periods of increasing errors in the seasonal cycles of Figure 10).

4. Conclusions

The analog forecast system presented here has been shown to be competitive with state-of-the-art dynamical forecast models in the Arctic despite its underperformance relative to dynamical models in middle latitudes. This regional dependence of the relative accuracies of the different approaches may be surprising at first glance, but it is consistent with the sparser network of *in situ* observations in the Arctic. Especially over the Arctic Ocean, which is the domain of the comparison presented here, there are no rawinsondes or surface observing stations.



Figure 10. Root-Mean-Square Errors (°C) of one-month temperature forecasts produced by the analog system (blue), the NMME (orange) and the CFSv2 model (green).

Aircraft and ship reports are also much less frequent in the Arctic, where surface buoys and satellite products are the main sources for the initialization of dynamical models. While one may argue that an analog forecast should also be sensitive to the initial state, the impact of uncertainties in the initial state is evidently less in an approach that inherently assumes some correspondence in the evolution of the atmosphere in years with similar initial states. Given the limits of deterministic predictability (one to two weeks at present), surface boundary forcing or some other "memory" in the system is being utilized by the analog system.

There are several notable limitations of the analog forecast system. First, the pool of candidate years for the analog matches is clearly limited, in this case to approximately 70 years. An analog-based forecast would clearly benefit from a larger sample of candidate analog years, as the closeness of the best match(es) will generally increase as the size of the candidate matches increases. Stated differently, the divergence from the "best" analogs will be more rapid as the closeness of the match worsens. Another major limitation is that the effect of climate change or even low-frequency internal climate variability. If climate change of variability takes the climate system to a new extreme that is outside the range spanned by the available analogs, then not only is a perfect match impossible but a close match becomes increasingly unlikely. The fact that the Arctic climate is changing rapidly and reaching new extremes (e.g., of temperature) makes this

limitation especially important in the Arctic.

The design of a system in which the predictor domain as well as the predictor variables can be varied makes the analog system presented here a natural candidate for machine learning. Machine learning has indeed been introduced into analog usage for short-range (days) weather forecasting [11]. With regard to the various user options included here, the combination of predictor regions, predictor variables, and predictor lead times that fare best in terms of forecast accuracy will vary, making optimization computationally challenging. However, if the system can "learn" from its successes and failures, it can move towards self-selection of optimal combinations of input parameters for a particular forecast. While such an approach will require additional levels of complexity in the system design, it could ultimately enable the analog method to realize its full potential in monthly and seasonal forecasting.

Funding

This work was supported by the National Oceanic and Atmospheric Administration's Climate Program Office through Grants NA15OAR4310169 and NA16-OAR4310162 as well as the International Arctic Research Center's (IARC) Experimental Arctic Prediction Initiative (EAPI).

Acknowledgements

The current version of the software was developed by the Scenarios Network for Alaska and Arctic Planning (B. Crevensten. K. Redilla, R. Torgerson, M. DeLue, C. Rosner).

Authors' Contributions

John Walsh, Nathan Kettle and Brian Brettschneider prepared the manuscript, with input from Richard Thoman. The project was supervised by John Walsh and Richard Thoman. Brian Brettschneider of the Alaska Center for Climate Assessment and Policy (now with the National Weather Service, Anchorage) developed the science and code for this tool

Data Availability Statement

The data used in generating the analog forecasts is the NCEP/NCAR R1 reanalysis (NCEP = National Centers for Environmental Prediction, NCAR = National Centers for Atmospheric Research). The reanalysis data are available for downloading at <u>https://psl.noaa.gov/data/gridded/data.ncep.reanalysis.pressure.html</u>. The monthly fields used here can be visualized using the online display tool of NOAA's Earth System Research:

https://psl.noaa.gov/cgi-bin/data/composites/printpage.pl.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- Lorenz, E.N. (1969) Atmospheric Predictability as Revealed by Naturally Occurring Analogues. *Journal of the Atmospheric Sciences*, 26, 636-646. <u>https://doi.org/10.1175/1520-0469(1969)26<636:APARBN>2.0.CO;2</u>
- Bergen, R.E. and Harnack, R.P. (1982) Long-Range Temperature Prediction Using a Simple Analog Approach. *Monthly Weather Review*, **110**, 1082-1099. <u>https://doi.org/10.1175/1520-0493(1982)110<1083:LRTPUA>2.0.CO;2</u>
- Gutzler, D.S. and Shukla, J. (1984) Analogs in the Wintertime 500 mb Height Field. Journal of the Atmospheric Sciences, 41, 177-189. https://doi.org/10.1175/1520-0469(1984)041<0177:AITWMH>2.0.CO;2
- Toth, Z. (1988) Long-Range Weather Forecasting Using an Analog Approach. *Journal of Climate*, 2, 594-607. https://doi.org/10.1175/1520-0442(1989)002<0594:LRWFUA>2.0.CO;2
- [5] Van den Dool, H. (1989) A New Look at Weather Forecasting through Analogues. *Monthly Weather Review*, 117, 2230-2247. <u>https://doi.org/10.1175/1520-0493(1989)117<2230;ANLAWF>2.0.CO;2</u>
- [6] NOAA (2018) El Nino Global Impacts. https://www.climate.gov/enso
- [7] Monache, L.D., Eckel, F.A., Rife, D.L., Nagarajan, B. and Searight, L. (2013) Probabilistic Weather Prediction with an Analog Ensemble. *Monthly Weather Review*, 141, 3498-3516. <u>https://doi.org/10.1175/MWR-D-12-00281.1</u>
- [8] Sperati, S., Alessandrini, S. and Monache, L.D. (2017) Gridded Probabilistic Weather Forecasts with an Analog Ensemble. *Quarterly Journal of the Royal Meteorological Society*, **143**, 2874-2885. <u>https://doi.org/10.1002/qi.3137</u>
- [9] Yu, H., Huang, J. and Chou, J. (2014) Improvement of Medium-Range Forecasts Using the Analog-Dynamical Method. *Monthly Weather Review*, 142, 1570-1587. https://doi.org/10.1175/MWR-D-13-00250.1
- [10] Eckel, F.A. and Monache, L.D. (2015) A Hybrid NWP-Analog Ensemble. Monthly Weather Review, 144, 897-911. <u>https://doi.org/10.1175/MWR-D-15-0096.1</u>
- [11] Chattopadhyay, A., Nabizadeh, E. and Hassanzadeh, P. (2020) Analog Forecasting of Extreme-Causing Weather Patterns Using Deep Learning. *Journal of Advances in Modeling Earth Systems*, 12, e2019MS001958. https://doi.org/10.1029/2019MS001958
- [12] Sima, R.J. (2020) Combining AI and Analog Forecasting to Predict Extreme Weather. *Eos*, 101. <u>https://doi.org/10.1029/2020EO140896</u>
- Kalnay, E., et al. (1996) The NCEP/NCAR Reanalysis 40-Year Project. Bulletin of the American Meteorological Society, 77, 437-471. https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2
- Kistler, R., *et al.* (2001) The NCEP-NCAR 50-Year Reanalysis: Monthly Means CD-ROM and Documentation. *Bulletin of the American Meteorological Society*, 82, 247-267. <u>https://doi.org/10.1175/1520-0477(2001)082<0247:TNNYRM>2.3.CO;2</u>
- [15] Walsh, J.E. and Brettschneider, B. (2019) Attribution of Recent Warming in Alaska. *Polar Science*, 21, 101-109. <u>https://doi.org/10.1016/j.polar.2018.09.002</u>
- Tippett, M.K. and DelSole, T. (2013) Constructed Analogs and Linear Regression. *Monthly Weather Review*, 141, 2519-2525. <u>https://doi.org/10.1175/MWR-D-12-00223.1</u>
- [17] Shukla, S., Funk, C. and Hoell, A. (2014) Using Constructed Analogs to Improve Skill of National Multi-Model Ensemble March-April-May Precipitation Forecasts

in Equatorial East Africa. *Environmental Research Letters*, **9**, Article ID: 094009. https://doi.org/10.1088/1748-9326/9/9/094009

- [18] Frediani, M.E., Hopson, T.M., Hacker, J.P., Anagnostou, E.N., Monache, L.D. and Vandenberghe, F. (2017) Object-Based Analog Forecasts for Surface Wind Speed. *Monthly Weather Review*, **145**, 5083-5102. https://doi.org/10.1175/MWR-D-17-0012.1
- [19] Van den Dool, H. (1994) Searching for Analogues, How Long Must We Wait? *Tellus*, **46**, 314-324. <u>https://doi.org/10.1034/j.1600-0870.1994.t01-2-00006.x</u>
- [20] NOAA (2020) Arctic Report Card: Update for 2020. Arctic Program, National Oceanic and Atmospheric Administration. <u>https://arctic.noaa.gov/Report-Card/Report-Card-2020</u>
- [21] Becker, E., Kirtman, B.P. and Pegion, K. (2020) Evolution of the North American Multi-Model Ensemble. *Geophysical Research Letters*, 47, e2020GL087408. <u>https://doi.org/10.1029/2020GL087408</u>
- [22] Saha, S., et al. (2014) The NCEP Climate Forecast System Version 2. Journal of Climate, 27, 2185-2208. <u>https://doi.org/10.1175/JCLI-D-12-00823.1</u>
- [23] Przybylak, R. (2020) Temporal and Spatial Variation of Surface Air Temperature over the Period of Instrumental Observations in the Arctic. *International Journal of Climatology*, 20, 587-614.
 https://doi.org/10.1002/(SICI)1097-0088(200005)20:6<587::AID-JOC480>3.0.CO;2-H