

RESEARCH ARTICLE

Data Analysis of Seasonal Air Temperature Variations in East Asia Using the Meteorological Datasets

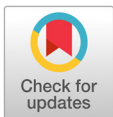
Jae-Hee Cho · Hak-Sung Kim*
Korea National University of Education

*Corresponding Author: envir007@knue.ac.kr

ABSTRACT

This study suggested to the National Center for Atmospheric Research (NCAR) command language (NCL) software setup in a Community Enterprise Operating System (CentOS) Linux-based computing environment, analyzing seasonal air temperature variations using the meteorological big data in the file format of the Network Common Data Form (NetCDF). The mathematic method of empirical orthogonal function (EOF), included as in the libraries within NCL, was used to analyze the statistical features of the East Asian atmospheric variability for the air temperatures at 850 hPa during the three climatological periods of 1971-2000, 1981-2010, and 1991-2020. The seasonal spatial patterns derived from EOF analysis showed that the seasonal air temperature variations in Korea, a downwind region in East Asia, were consistent with those in Mongolia and northern China, primarily throughout all seasons. The influence of the East Asian continent on Korea has gradually intensified during the early spring season. Furthermore, the expanding air mass over the Northwest Pacific to the East Asian region in summer has gradually intensified the influence on the air temperatures in Korea for the recent climatological period of 1991-2020. The dominant EOF mode is a monopole covering most of East Asia. The seasonal principal component analysis (PCA) on the first EOF mode in East Asia shows that it significantly correlates with the variations of the observed air temperatures in Korea ahead of the recent climatological period.

Key words: Meteorological big data, EOF analysis, air temperature seasonal variations, NCL, NetCDF, Linux



OPEN ACCESS

Brain, Digital, & Learning
2022, Vol. 12, No. 4, 637-647.
<https://doi.org/10.31216/BDL.20220039>

Received: November 11, 2022

Revised: November 27, 2022

Accepted: December 28, 2022

© 2022. Institute of Brain based Education,
Korea National University of Education



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

Big data refers to overwhelming amounts of complex and heterogeneous data to be dealt with by traditional data-processing applications. In recent years, data science has been the domain of study that deals with these big data using modern tools and techniques to find unseen patterns and derive meaningful information (Agarwal & Dhar, 2014; Dobre & Xhafa, 2014; Sivarajah et al., 2017). Big data in the Earth sciences are conducted through the development of nascent technological instruments such as meteorological satellites, ocean observing systems, seismic instrumentation systems, and space telescopes (Gvishiani et al., 2022). A vast amount of data from the core of the Earth to the universe are analyzed using advanced computational scientific methods. In atmospheric science, modern findings have been conducted by developing instruments and numerical modeling using the high-performance computing system on broad spatiotemporal scales (Fathi et al., 2021; Zhou et al., 2022). On the other hand, storage of the weather and climate datasets is necessary for complex multidimensional file formats of big sizes. Network Common Data Form (NetCDF) is the file format storing multidimensional scientific variables and is widely adopted in Earth science research such as meteorology, climatology, geology, and astronomy. Traditional scientific languages such as C, C++, and Fortran can read and manipulate NetCDF files. However, NetCDF programming interfaces are more readily available in dealing with modern languages of Python, R, Interface Description Language (IDL), MATrix LABoratory (MATLAB), and NCAR Command Language (NCL).

An analysis of the meteorological data is conducted to objectively analyze weather and climate, identify mechanisms, and determine the representative data for a region of interest in a globe. With the availability of vast amounts of meteorological data, the demand for techniques and tools for analyzing them also increased in recent years (Fathi et al., 2021; Zhou et al., 2022). The NCL, a product of the National Center for Atmospheric Research (NCAR) sponsored by the National Science Foundation in the United States of America, is a free interpreted language for data processing and visualization. NCL supports a variety of file formats, including NetCDF, for use in meteorological research. Currently, NCL is no longer updated and the transition to Python module form is in full swing. Recent trends show that open-source languages are becoming highly popular because they support major functionalities required for meteorological data analysis and parallel computing (Gharat et al., 2022).

Meanwhile, East Asia is undergoing significant climate changes from a wide range of factors including extreme temperature and precipitation, and it is likely to grow (Cho et al., 2021; Cho et al., 2022; Kim et al., 2016). In particular, the wind speeds near the tropopause tend to weaken as the meridional air temperature gradient decreases due to the intensified atmospheric warming in the higher latitudes of East Asia (Kim et al., 2016). Such wind speed changes in East Asia have resulted in changing atmospheric environments in Korea (Cho et al., 2021; Cho et al., 2022). The NCL could be used to analyze the variations of climate normal, indicating the 30-year average of meteorological variables in East Asia. Moreover, the seasonal climatological mean in East Asia could be described using massive meteorological data. This study emphasizes interpreting meteorological big data and analyzing seasonal air temperature variations in East Asia during the three climatological periods of 1971-2000, 1981-2010, and 1991-2020.

Materials and Methods

Meteorological Reanalysis Data

A reanalysis is a systematic approach to producing meteorological data sets on multidimensional scales. Meteorological reanalysis data incorporate millions of observations into a stable data assimilation system, enabling an analysis of weather and climate processes (Kalnay et al., 1996). Reanalysis data sets are multivariate, spatially, temporally complete, and gridded. Also, meteorological reanalysis data combine incomplete and inaccurate observations and imperfect models.

The National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) reanalysis dataset for the meteorological variables consists of atmospheric model outputs with different time scales of 4-times daily, the daily mean, and monthly mean from the year 1948 up to the present. The monthly air temperatures of NCEP/NCAR reanalysis datasets (<https://psl.noaa.gov/data/>) were taken to analyze atmospheric variability in East Asia using NCL libraries, which datasets are available in the NetCDF file format. Fig. 1 shows the air temperature NetCDF file format, indicating the detailed structure of the four dimensions. The NCEP/NCAR air temperature datasets comprise four dimensions of time, pressure level, latitude, and longitude; monthly intervals from 1948 to the present, pressure levels from 1,000 hPa to 10 hPa, and spatial coverage of 2.5° latitudes $\times$ 2.5° longitude. In this study, the monthly air temperatures were used in spatial coverage of East Asia with $25^\circ\text{--}55^\circ\text{N}$ and $90^\circ\text{--}150^\circ\text{E}$ at 850 hPa for 1971–2020.

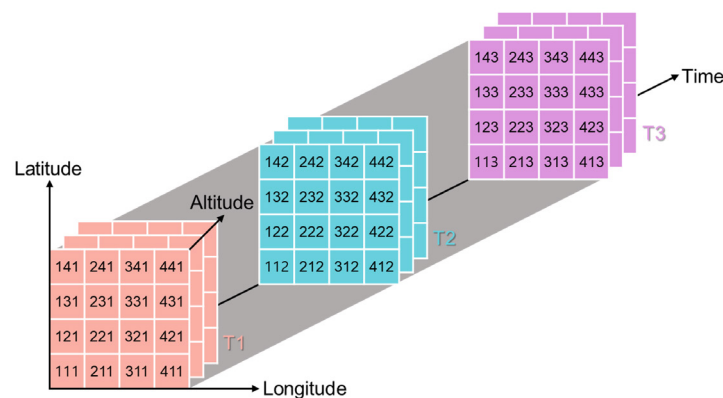


Fig. 1. Meteorological dataset as in a NetCDF file format comprising four-dimension scales

Empirical Orthogonal Function Analysis

The most widely used statistical technique in meteorology is principal component analysis (PCA). The PCA technique has gained ground for studying meteorological data by Lorenz (1956), who called the technique an empirical orthogonal function (EOF) analysis. The EOF analysis finds the spatial and temporal variability and provides the importance of each pattern in meteorological studies. An EOF analysis is a technique that mathematically compresses N -dimensional data and expresses it in a dimension smaller than N . The analysis data can be linearly transformed to extract the significant variability from the calculated data, which can explain most of the raw data. The columns of matrix U are called the EOF

of A , and the diagonal elements of matrix S are called eigenvalues. The eigenvalues of the matrix S are the variance of the raw data described by EOF, and the sum of the eigenvalues of all EOFs gives the variance of the raw data in the following Eq. 1. Each column of the matrix V^T represents the temporal evolution of each EOF as a time series coefficient.

$$A = U \times S \times V^T \text{ (Eq. 1)}$$

In this study, an EOF analysis was employed to analyze the impacts of the East Asian seasonal air temperature variations on Korea. The area targeted for the EOF analysis was consistent with the coverage of the NCEP/NCAR monthly air temperature dataset in East Asia (90-150°E, 25-55°N), including eastern China and Korea. To analyze the seasonal air temperature variations depending on seasons between the three climatological periods of 1971-2000, 1981-2010, and 1991-2020, an EOF analysis was applied to the monthly air temperatures at 850 hPa. EOF analysis was performed using the seasonal average value of monthly air temperature. The principal components were calculated in the first and second modes in the EOF analysis. Then, the seasonal PCA results were compared with surface-based air temperature observations from the Korea Meteorological Administration (KMA) for the three climatological periods, describing the effects of the East Asian continent on Korea.

Results and Discussions

Setup of the NCL in the Linux-based System

NCL is a powerful modern language for reading, writing, manipulating, and visualizing meteorological NetCDF file datasets. In this study, NCL, version 6.6.2, was installed on the 64-bit Community Enterprise Operating System (CentOS) environment, comprising a hardware system of CPU with 64 cores (2.7 GHz) and memory (1TB). CentOS, version 7.9, was already installed in the hardware system. In this study, NCL was built with the GNU compilers in the CentOS Linux-based system. A CentOS Linux-based NCL was efficiently used through the SSH-connecting terminal software installed in the Windows operating system-based computer (Windows system). The SSH connections from the Windows system enabled it to command performing NCL in a script file in a CentOS Linux-based system and to get the analysis results within a postscript file format. NCL contains libraries for a suite of functions to read NetCDF meteorological data files, analyze EOF, and plot analysis results.

Every 10 years, average values for air temperatures, rainfalls, and other meteorological variables are calculated for the past three decades, known as the climate normal; those 30-year averages now span 1991-2020. Fig. 2 shows the NCL script reading and plotting NCEP/NCAR monthly air temperatures to analyze EOF spatial patterns during the three climatological periods of 1971-2000, 1981-2010, and 1991-2020. The parameters specifying the region and period are defined, and the data corresponding to the specified region and period is open. The standard EOF function, included in the NCL libraries indicating mathematical Eq. 1, was declared in the script to conduct spatial patterns and the seasonal PCA annual series. The EOF analysis results are produced within a plot in the workstation under the Windows system and converted the analysis results to a postscript file format. Those EOF functions would allow the NCL libraries to become educational materials in Earth science coursework and help instructors provide students with a reason to learn scientific languages.

```

1 begin
2 ;=====
3 ; User defined parameters that specify region of globe and
4 ;=====
5 latS = 25.
6 latN = 55.
7 lonL = 90.
8 lonR = 150.
9
10 yrStrt = 1971 ; 1981 ; 1991
11 yrLast = 2000 ; 2010 ; 2020
12
13 season = "DJF"; "HAM"; "JJA"; "SON" ; choose seasonal mean
14
15 neof = 2 ; number of EOFs
16 optEOF = True
17 optEOF@opt = 0 ; This is the default; most commonly used; no need to specify.
18
19 optETS = False
20
21 ;=====
22 ; Open the file: Read only the user specified period
23 ;=====
24 f = addfile("air.mon.mean.nc", "r")
25
26 TIME = f->time
27 YYYY = cd_calendar(TIME, -1)/100 ; entire file
28 YYYY = ind(YYYY.ge.yrStrt .and. YYYY.le.yrLast)
29
30 slp = f->air(iYYYY,(850),:,:)
31 printVarSummary(slp) ; variable overview
32
33 ;=====
34 ; dataset longitudes span @=357.5
35 ; Because EOFs of the North Atlantic Oscillation are desired
36 ; use the "lonFlip" (contributed.ncl) to reorder
37 ; longitudes to span -180 to 177.5: facilitate coordinate subscripting
38
39 slp = lonFlip(slp)
40 printVarSummary(slp) ; note the longitude coord
41
42 ;=====
43 ; compute desired global seasonal mean: month to season (contributed.ncl)
44 ;=====
45 SLP = month_to_season(slp, season)
46 nyrs = dimsizes(SLP@time)
47 printVarSummary(SLP)
48
49 ;=====
50 ; create weights: sqrt(cos(lat)) [or sqrt(gw)]
51 ;=====
52 rad = 4.*atan(1.)/180.
53 clat = f->lat
54 clat = sqrt(cos(rad*cLat)) ; gw for gaussian grid
55
56 ;=====
57 ; weight all observations
58 ;=====
59 wSLP = SLP
60 wSLP = SLP*conform(SLP, clat, 1) ; copy meta data
61 wSLP@long_name = "Wgt: "+wSLP@long_name
62
63 ;=====
64 ; Reorder (lat,lon,time) the "weighted" input data
65 ; Access the area of interest via coordinate subscripting
66 ;=====
67 xw = wSLP({lat|latS:latN},{lon|lonL:lonR},time,:)
68 x = wSLP(time::,{lat|latS:latN},{lon|lonL:lonR})
69
70 eof = eofunc.Wrap(xw, neof, optEOF)
71 eof_ts = eofunc_ts.Wrap(xw, eof, optETS)
72
73 printVarSummary(eof) ; examine EOF variables
74 printVarSummary(eof_ts)
75
76 ;=====
77 ; Normalize time series: Sum spatial weights over the area of used
78 ;=====
79 dimxw = dimsizes(xw)
80 min = dimxw(1)
81 sumWgt = min*sum(clat({lat|latS:latN}))
82 eof_ts = eof_ts/sumWgt
83
84 ;=====
85 ; Extract the YYYYMM from the time coordinate
86 ; associated with eof_ts (same as SLP@time)
87 ;=====
88
89 yyyyymm = cd_calendar(eof_ts@time, -2)/100

```

```

90 ;=====
91 ; PLOTS
92 ;=====
93
94 wks = gsn_open_wks("ps", "eof") ; send graphics to PostScript file
95 plot = new(neof,graphic) ; create graphic array
96 ; only needed if paneling
97
98 ; EOF patterns
99
100 res = gsnDraw = True ; don't draw yet
101 res@gsnFrame = False ; don't advance frame yet
102
103 res@gsnAddCyclic = False ; plotted data are not cyclic
104
105 res@mpFillon = False ; turn off map fill
106 res@mpMinLatF = latS ; zoom in on map
107 res@mpMaxLatF = latN
108 res@mpMinLonF = lonL
109 res@mpMaxLonF = lonR
110
111 res@cnFillon = True ; turn on color fill
112 res@cnLinesOn = False ; True is default
113 res@cnFillPalette = "dLWbRte" ; set color map
114 res@lbLabelBarOn = False ; turn off individual lb's
115
116 symMinMaxPlt(eof, 16, False, res) ; set symmetric plot min/max
117 ; contributed.ncl
118
119 ; panel plot only resources
120 resP = gsnMaximize = True ; modify the panel plot
121 resP@gsnMaximize = True ; large format
122 resP@gsnPanelLabelBar = True ; add common colorbar
123
124 yStrt = yyyyymm(0)/100
125 yLast = yyyyymm(nyrs-1)/100
126 resP@gsnPanelMainString = "SLP: "+season+": "+yStrt+"-"+yLast
127
128 ;=====
129 ; first plot
130 ;=====
131 do n=0,neof-1
132   res@gsnLeftString = "EOF "+(n+1)
133   res@gsnRightString = sprintf("%5.1f", eof@pcvar(n))+"%"
134   plot(n)=gsn_csm_contour_map(wks,eof(n,:),res)
135 end do
136 gsn_panel(wks,plot,(/neof,1/),resP) ; now draw as one plot
137
138 ;=====
139 ; second plot
140 ;=====
141 ; EOF time series [bar form]
142
143 rts = gsnDraw = True ; don't draw yet
144 rts@gsnFrame = False ; don't advance frame yet
145 rts@gsnScale = True ; force text scaling
146
147 ; these four resources allow the user to stretch the plot size, and
148 ; decide exactly where on the page to draw it.
149
150 rts@vpHeightF = 0.40 ; Changes the aspect ratio
151 rts@vpWidthF = 0.85
152 rts@vpXF = 0.10 ; change start locations
153 rts@vpYF = 0.75 ; the plot
154
155 rts@tiYAxisString = "Pa" ; y-axis label
156
157 rts@gsnYRefLine = 0. ; reference line
158 rts@gsnYBarChart = True ; create bar chart
159 rts@gsnAboveYRefLineColor = "red" ; above ref line fill red
160 rts@gsnBelowYRefLineColor = "blue" ; below ref line fill blue
161
162 ;=====
163 ; panel plot only resources
164 ;=====
165 rtsP = gsnMaximize = True ; modify the panel plot
166 rtsP@gsnMaximize = True ; large format
167 rtsP@gsnPanelMainString = "SLP: "+season+": "+yStrt+"-"+yLast
168
169 year = yyyyymm/100
170 ; create individual plots
171 do n=0,neof-1
172   rts@gsnLeftString = "EOF "+(n+1)
173   rts@gsnRightString = sprintf("%5.1f", eof@pcvar(n))+"%"
174   plot(n) = gsn_csm_xy(wks,year,eof_ts(n,:),rts)
175 end do
176 gsn_panel(wks,plot,(/neof,1/),rtsP) ; now draw as one plot
177 end

```

Fig. 2. A script of NCL calculating seasonal spatial patterns and PCAs in the EOF modes for monthly air temperatures over the East Asian region

Seasonal Air Temperature Variations in East Asia

The climate normal variations for seasonal air temperatures in East Asia were analyzed using the EOF method. Fig. 3 shows the spatial EOF patterns of the seasonal air temperatures at 850 hPa in East Asia (90-150°E, 25-55°N) during each climatological period of 1971-2000, 1981-2010, and 1991-2020. The EOF modes are dominated as a monopole covering most of East Asia. While the seasonal air temperatures varied a little bit among the three climatological periods, the first EOF mode for the most recent climatological period of 1991-2020 accounted for 50.6% in spring, 50.3% in winter, 42.7% in fall, and 36.9% in summer as in the same manner in the East Asian region ranging from northern Mongolia to Korea. Furthermore, the second EOF mode accounted for 25.7% of fall, 18.3% of summer, 17.5% of spring, and 16.2%

of winter, with dipole structures of differing signals between northeastern and southwestern Asia. In other studies, on East Asia, atmospheric circulation anomalies have essential contributions to the formation of seasonal meteorological variable anomalies in East Asia (Chen et al., 2016; Chen et al., 2019; Miyazaki & Yasunari, 2008).

During the recent climatological period of 1991-2020, the sum of the two EOFs accounts for 68.4% in fall, 68.1% in spring, 67.5% in winter, and 55.2% in summer. Furthermore, the seasonal EOF percentages showed concurrent air temperature variations in the East Asian region from Mongolia eastward to Korea through eastern China during the three climatological periods. The East Asian continent has influenced Korea by the prevailing westerlies. Kim et al. (2017) reported that significant increases in annual air temperatures grew more substantial in the higher latitudes over 40°N such as Mongolia and northeastern China, and the northwest Pacific Ocean, compared with the mid-latitude region of eastern China and Korea from 1997 to 2016.

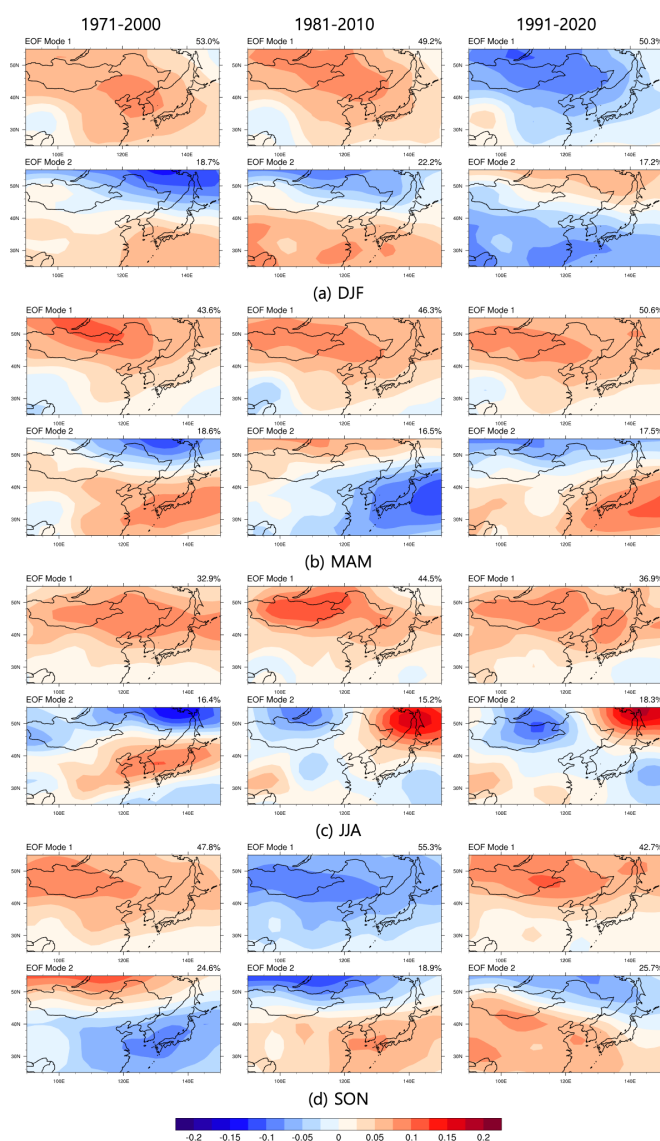


Fig. 3. Seasonal spatial patterns of the EOF modes for the monthly air temperatures at 850 hPa regarding (a) DJF, (b) MAM, (c) JJA, and (d) SON in the East Asian region during the three climatological periods of 1971-2000, 1981-2010, and 1991-2020.

Fig. 4 shows the time series of the EOF modes for the seasonal air temperatures at 850 hPa in East Asia (90–150°E, 25–55°N) during each climatological period of 1971–2000, 1981–2010, and 1991–2020. In the seasonal air temperature variations, the influence of the East Asian continent on Korea has declined with lowering atmospheric warming effects in winter during the recent 1991–2020 than those previous climatological periods. A few studies suggested decreases in the air temperature increasing rate during the winter season in East Asia (Min et al., 2015; You et al., 2022). In contrast, the air temperature increasing rate has gradually intensified during the warm seasons compared with the winter seasons. While the weather conditions of the Korean summer have been dominated by the expanding air mass of the Northwest Pacific, the influence of the East Asian continent on Korea gradually intensified during spring and summer during the recent climatological period of 1991–2020.

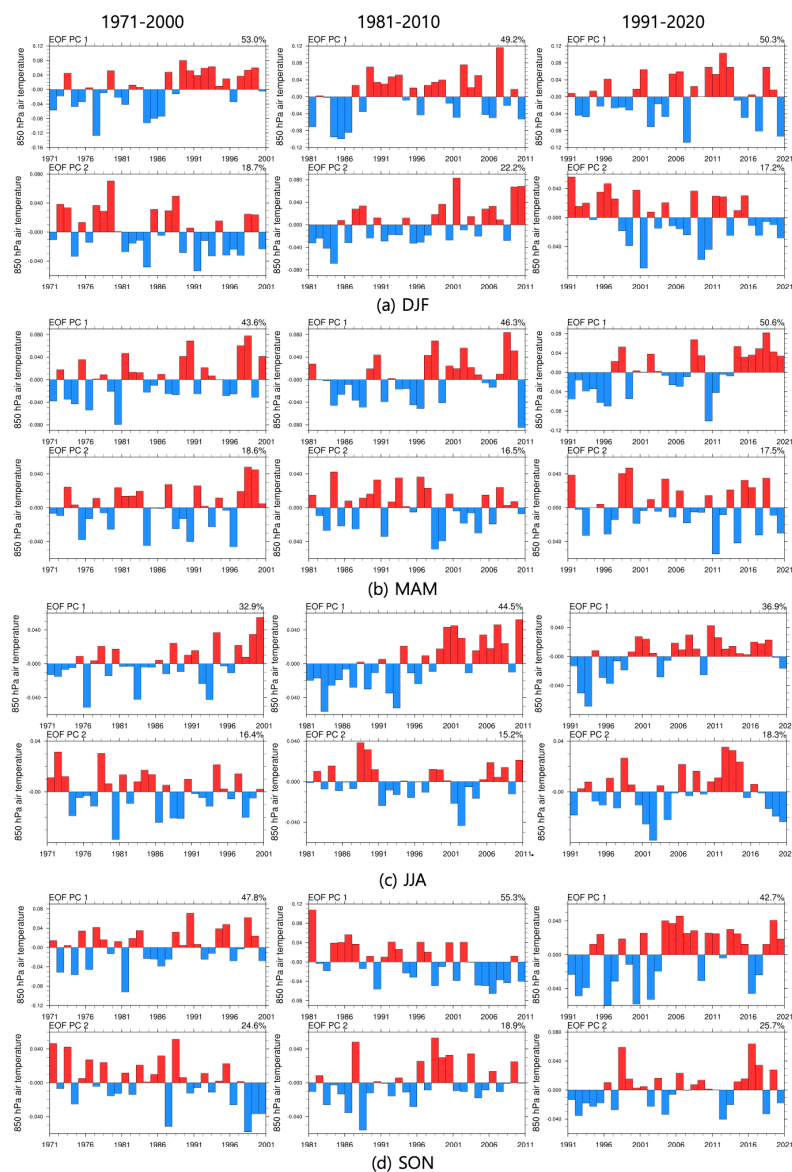


Fig. 4. Seasonal time-series of the EOF modes for the monthly air temperatures at 850 hPa regarding (a) DJF, (b) MAM, (c) JJA, and (d) SON in the East Asian region during the three climatological periods of 1971–2000, 1981–2010, and 1991–2020.

Impact of Seasonal Variations in East Asia on Korea

Seasonal PCA on the first EOF modes in East Asia was compared with the surface-based observations of air temperatures in Korea to analyze the variations of the impacts of the East Asian continent on Korea during the climatological periods. The surface-based observations of air temperatures in Korea were analyzed using the national average values from the Korea Meteorological Administration during 1971-2020. Since the first EOF mode contains more than twice the second mode, the first EOF mode is mainly analyzed. Fig. 5 shows the correlation coefficients in seasons between the PCAs and the observed air temperatures in Korea during the climatological periods from 1971-2000 to 1991-2020. The seasonal PCAs in East Asia were correlated with air temperatures in Korea on significant levels of less than 0.05 in spring, summer, and fall. The correlation coefficient gradually increased during spring and summer for the climatological periods from 1971-2000 to 1991-2020.

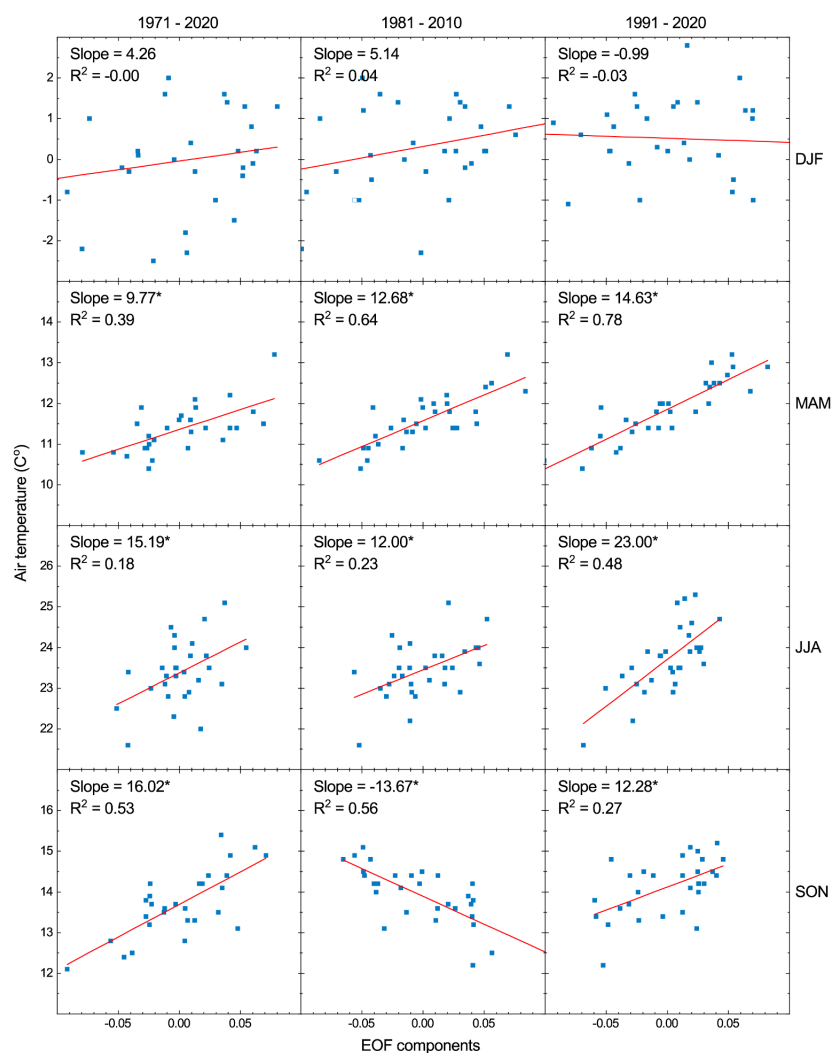


Fig. 5. Seasonal correlation coefficients between PCA results on the first EOF modes in East Asia and the surface-based air temperatures in Korea for the three climatological periods of 1971-2000, 1981-2010, and 1991-2020. (*The slope represents statistical significance at a 95% confidence level)

The highest level of the correlation coefficients exceeding 0.78 ($p < 0.05$) was recorded in the spring of the recent climatological period of 1991-2020. The increasing trend in air temperatures was found to have intensified in the spring season, compared with those in other seasons in East Asia (Cho et al., 2022; Jung et al., 2002; Kim et al., 2017; Kug & Ahn, 2013). Especially, Cho et al. (2021) showed a recently intensified increasing trend of the air temperature during spring in East Asia through an analysis of air temperature anomalies throughout 1998-2019. Moreover, in 2018, unusual East Asian spring warming occurred, leading to a North Atlantic tripole SST-mode transportation anomaly (Deng et al., 2019). The correlation coefficients in summer explained that the impact of Northwest Pacific air mass had intensified variations of the air temperatures in Korea ahead of the recent climatological period. In addition, Lee & Lee (2016) suggested that the summer duration had been extended in Korea for the past 42 years (1973-2014). The anomalous high-pressure pattern was accompanied by large-scale subsidence in Korea, providing favorable conditions for sweltering and humid days.

Conclusions

This study presented an NCL software setup in a CentOS Linux-based hardware system. It analyzed the impacts of East Asian atmospheric variability on Korea located on the downwind side during the climatological periods of 1971-2000, 1981-2010, and 1991-2020. In the meteorological data analysis domain, the NCL tool is widely adopted to analyze the big data file format as in NetCDF. The open-source availability of software and advanced scientific programming concepts, and extensive community support provide an additional advantage in using it. NCL routines were successfully installed on the CentOS Linux-based hardware system, which conducted the spatial pattern and PCA results using the EOF analysis method for the monthly air temperatures at 850 hPa in East Asia.

While the seasonal air temperatures varied among the climatological periods, the first EOF mode accounted for 50.6% in spring, 50.3% in winter, 42.7% in fall, and 36.9% in summer in the same signal from northern Mongolia to Korea during the recent climatological period of 1991-2020. The EOF analysis represented a spatial structure in which the variations were consistent with those in Korea during all seasons. The East Asian continent has influenced the Korean weather by the prevailing westerlies. In the seasonal air temperature variations, the impact of the East Asian continent on Korea in winter has decreased more during the recent 1991-2020 than during the previous climatological periods. In particular, the air temperatures in the East Asian continent have gradually intensified in Korea during the warm season of spring and summer. The air mass over the Northwest Pacific has strengthened the variability of air temperatures in Korea in summer for the recent climatological period of 1991-2020. Seasonal PCA on the first EOF modes in East Asia was compared with the air temperatures observed in Korea for the climatological periods. The correlation coefficients between PCAs in the East Asian region and the surface-based air temperatures in Korea indicated gradual increases in spring and summer ahead of the recent climatological period. In addition, the expanding air mass over the Northwest Pacific to the East Asian region in summer has intensified the dominant effects on the air temperatures in Korea.

References

- Agarwal, R. & Dhar, V. (2014). Editorial – big data, data science, and analytics: The opportunity and challenge for is research. *Information Systems Research*, 25, 443-448.
- Chen, W., Hong, X., Lu, R., Jin, A., Jin, S., Nam, J. C., Shin, J. H., Goo, T. Y., & Kim, B. J. (2016). Variation in summer surface air temperature over Northeast Asia and its associated circulation anomalies. *Advances in Atmospheric Sciences*, 33, 1-9.
- Chen, S., Wu, R., Song, L., & Chen, W. (2019). Interannual variability of surface air temperature over mid-high latitudes of Eurasia during boreal autumn. *Climate Dynamics*, 53, 1805-1821.
- Cho, J. H., Kim, H. S., & Chung, Y. S. (2021). Spatio-temporal changes of PM10 trends in South Korea caused by East Asian atmospheric variability. *Air Quality, Atmosphere & Health*, 14, 1001-1016.
- Cho, J. H., Kim, H. S., & Yoon, M. B. (2022). The influence of atmospheric blocking on regional PM10 aerosol transport to South Korea during February–March of 2019. *Atmospheric Environment*, 277, 119056.
- Deng, K., Yang, S., Lin, A., Li, C., & Hu, C. (2019). Unprecedented East Asian warming in spring 2018 linked to the North Atlantic tripole SST mode. *Atmospheric and Oceanic Science Letters*, 12, 246-253.
- Dobre, C. & Xhafa, F. (2014). Intelligent services for big data science. *Future Generation Computer Systems*, 37, 267-281.
- Fathi, M., Haghi Kashani, M., Jameii, S. M., & Mahdipour, E. (2021). Big data analytics in weather forecasting: A systematic review. *Archives of Computational Methods in Engineering*, 29, 1247–1275.
- Gharat, J., Kumar, B., Ragha, L., Barve, A., Jeelani, S. M., & Clyne, J. (2022). Development of NCL equivalent serial and parallel python routines for meteorological data analysis. *The International Journal of High Performance Computing Applications*, 36, 337-355.
- Gvishiani, A. D., Dobrovolsky, M. N., Dzeranov, B. V., & Dzeboev, B. A. (2022). Big data in geophysics and other earth sciences. *Izvestiya, Physics of the Solid Earth*, 58, 1-29.
- Hulme, M., Zhao, Z. C., & Jiang, T. (1994). Recent and future climate change in East Asia. *International Journal of Climatology*, 14, 637-658.
- Jung, H. S., Choi, Y., Oh, J. H., & Lim, G. H. (2002). Recent trend in temperature and precipitation over Korea. *International Journal of Climatology*, 22, 1327-1337.
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., ... Joseph, D. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, 77, 437-472.
- Kim, H. S., Chung, Y. S., & Cho, J. H. (2017). Long-term variations of dust storms and associated dustfall and related climate factors in Korea during 1997-2016. *Air Quality, Atmosphere & Health*, 50, 191-200.
- Kim, H. S., Chung, Y. S., Tans, P. P., & Yoon, M. B. (2016). Climatological variability of air temperature and precipitation observed in South Korea for the last 50 years. *Air Quality, Atmosphere & Health*, 9, 645-651.
- Kug, J. S. & Ahn, M. S. (2013). Impact of urbanization on recent temperature and precipitation trends in the Korean peninsula. *Asia-Pacific Journal of Atmospheric Sciences*, 49, 151-159.
- Lee, W. S. & Lee, M. I. (2016). Interannual variability of heat waves in South Korea and their connection with large-scale atmospheric circulation patterns. *International Journal of Climatology*, 36, 4815-4830.
- Lorenz, E. N. (1956). Empirical orthogonal functions and statistical weather prediction. *Science Report 1, Statistical Forecasting Project, Department of Meteorology, MIT (NTIS AD 110268)*, 49 pp.
- Min, S. K., Son, S. W., Seo, K. H., Kug, J. S., An, S. I., Choi, Y. S., ... Lee, M. I. (2015). Changes in weather and climate extremes over Korea and possible causes: A review. *Asia-Pacific Journal of Atmospheric Sciences*, 51, 103-121.

- Miyazaki, C., & Yasunari, T. (2008). Dominant interannual and decadal variability of winter surface air temperature over Asia and the surrounding oceans. *Journal of Climate*, 21, 1371-1386.
- Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of big data challenges and analytical methods. *Journal of Business Research*, 70, 263-286.
- You, Q., Jiang, Z., Yue, X., Guo, W., Liu, Y., Cao, J., ... Zhai, P. (2022). Recent frontiers of climate changes in East Asia at global warming of 1.5°C and 2°C. *npj Climate and Atmospheric Science*, 5, 80.
- Zhou, H., Ren, H., Royer, P., Hou, H., & Yu, X.-Y. (2022). Big data analytics for long-term meteorological observations at Hanford site. *Atmosphere*, 13, 136.

Authors Information

Cho, Jae-Hee: Korea National University of Education, Researcher, First Author

ORCID: <https://orcid.org/0000-0002-4280-5267>

Kim, Hak-Sung: Korea National University of Education, Professor, Corresponding Author

ORCID: <https://orcid.org/0000-0001-5367-5683>