

# Warner Internship for Scientific Enrichment (WISE)

## Educational Materials for the Analog Ensemble Technique

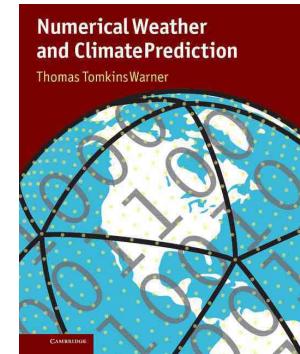
---



Laura Clemente-Harding  
7 July 2017  
Project Summary

# WISE Proposal

- Target audience
  - Upper level undergraduate students / Graduate students
- Develop approachable, concise teaching material on the AnEn technique including
  - Educational reading
  - Presentation slides for teaching the material
  - Lab: Hands on computing exercise
- Present this material to visiting undergraduate students



# Ties to graduate student ASP work

- Increase understanding of the point based AnEn technique
- Investigate a smart means of extending the search space for the point based AnEn technique
- Develop a spatially aware similarity metric for the AnEn technique

# Educational Materials

- Educational Document covers
  - Brief history of analogs related research in the weather community
  - How the AnEn technique developed
  - How the AnEn technique differs from previous analog related meteorological research
  - How the technique works
  - Overview of the current state of the art applications of the technique

## A Beginners Introduction to the Analog Ensemble Technique

Written by Laura Clemer  
Warner Internship for Sci  
Teaching Mater

### Sections

\*Please click on a section below to jump to the sections contents

1. [Introduction](#)
2. [Weather Analogs](#)
3. [Analog Ensemble Technique Overview](#)
4. [How the AnEn Technique Developed and How it Differs from Previous Analog Related Meteorological Research](#)
5. [Overview of Current State of the Art for the Analog Ensemble Technique](#)
6. [References](#)

### Appendix

1. [Acronyms](#)

**Point 1 (P1):**

- Determine the most similar past forecasts based on the similarity metric defined in [Delle Monache et al. \(2013\)](#):
- 4 ensemble members = (5, 2, 1, 10)
- Take the corresponding observation for each chosen forecast and identify the dates each observation came from:
  - Corresponding observations
  - Dates
- Dates = (00Z 12 Dec, 00Z 1 Dec, 00Z 10 Jan, 00Z 10 Mar)
- Rank the past similar forecasts and the observations
  - Ranked past forecasts = (1,2,5,10)
  - Ranked (sorted) observations = (2,4,9,10)
- Determine the B function (Copula function)
  - The B function is the connection between the ranked and unranked. This is what is used to shuffle the positions.
  - $B(P_{sorted \ obs}) = (1,2,3,4) = (2,1,4,3)$
  - The B function changes for every lead-time and every station.
- Now using this B function that links the position of the observation before sorting and the position after sorting, re-sort the ensemble forecast members. The result becomes:
  - (2, 1, 10, 5)
  - Position one (1) corresponds to the second (2<sup>nd</sup>) position. Position two (2) corresponds to the first (1<sup>st</sup>) position. Position three (3) corresponds to the fourth (4<sup>th</sup>) position in the vector. Position four (4) corresponds to the third (3<sup>rd</sup>) position in the vector.
- This is executed at lead-time = 0 (analysis). Next, the same process will be executed for lead-time = 1 and so on. However, the same dates (e.g. 12 Dec, 1 Dec, 10 Jan, 10 Mar) must not be used.

**Moving to the 2<sup>nd</sup> point:**

Note: (1) the selected dates have to remain the same for ensuing lead times and (2) the same dates have to be the same across multiple locations.

- Dates
- Dates = (00Z 12 Dec, 00Z 1 Dec, 00Z 10 Jan, 00Z 10 Mar)
- Same past forecast dates (could also call this members) but different forecast values and different observations because this is a different location
- 4 ensemble members = (6,1,2,9)

# Educational Materials

- Slides for teaching

- Lab Exercise

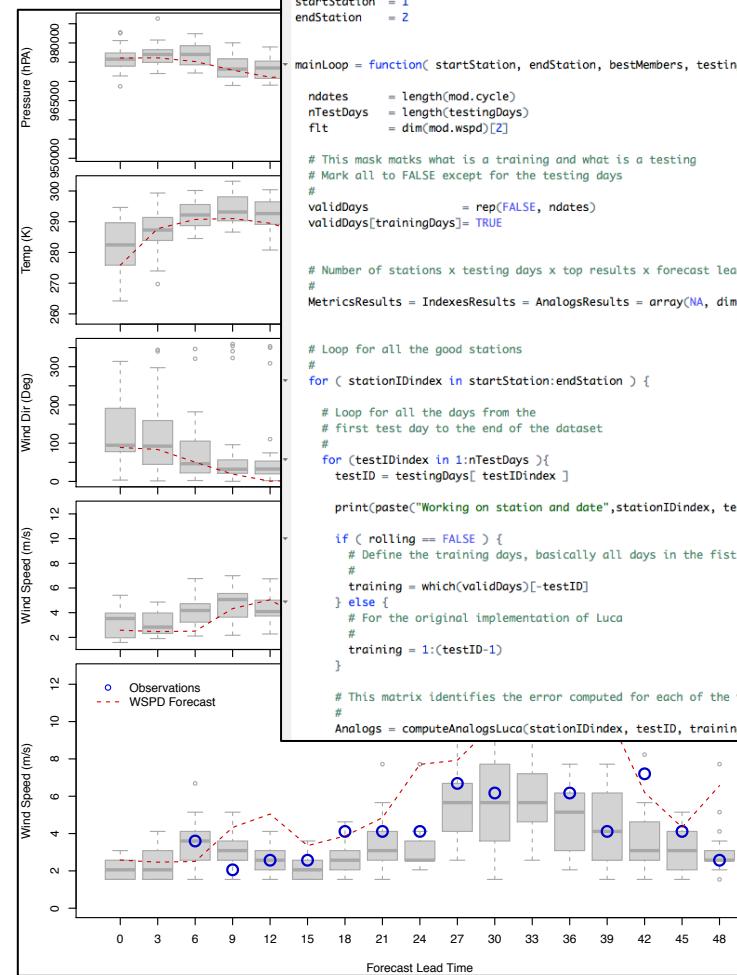
- Hands on computing exercise experiments which help to understand analog selection, number of ensemble members, and training size
- Stand alone R code that allows students to experiment with parameters

- Handout

- Summary of previous research and applications using the AnEn technique

## **A Beginners Introduction to the Analog Ensemble Technique:**

## Hands on Exercise with the Analog Ensemble Technique



```

bestMembers      = 110
trainingDays    = 1:357
testingDays     = 358:360 # The first year (leap year) 152:153#
rolling         = TRUE # For Luca's original implementation

startStation    = 1
endStation      = 2

mainLoop = function( startStation, endStation, bestMembers, testingDays, trainingDays, rolling)

  ndates      = length(mod.cycle)
  nTestDays   = length(testingDays)
  flt         = dim(mod.wspd)[2]

  # This mask marks what is a training and what is a testing
  # Mark all to FALSE except for the testing days
  #
  validDays      = rep(FALSE, ndates)
  validDays[trainingDays] = TRUE

  # Number of stations x testing days x top results x forecast lead times
  #
  MetricsResults = IndexesResults = AnalogsResults = array(NA, dim=c(endStation-startStation, nTestDays, bestMembers, 12))

  # Loop for all the good stations
  #
  for ( stationIDindex in startStation:endStation ) {

    # Loop for all the days from the
    # first test day to the end of the dataset
    #
    for (testIDindex in 1:nTestDays ) {
      testID = testingDays[ testIDindex ]

      print(paste("Working on station and date",stationIDindex, testIDindex))

      if ( rolling == FALSE ) {
        # Define the training days, basically all days in the first year, except for the current
        # test day
        training = which(validDays)[-testID]
      } else {
        # For the original implementation of Luca
        # training = 1:(testID-1)
      }

      # This matrix identifies the error computed for each of the variables
      #
      Analogs = computeAnalogsLuca(stationIDindex, testID, training )
    }
  }
}

```

# What's next ?

- RAL Seminar on the Analog Ensemble Technique
  - Tuesday, 25 July at 1pm (FL-2-001)  
Come learn about the AnEn!
- Obtain feedback from seminar, revise materials accordingly, provide revised version this Fall 2017 for Dr. Cervone's undergraduate course
- Fall 2017 Course Offerings at PSU will be using this material
  - Dr. George Young (undergraduate): Portions of this material will be used
  - Dr. Guido Cervone (undergraduate): This material will be expanded for a 2.5 hour class session

Thank you for the opportunity!

Special thanks to Dr. Sarah Tessendorf, Dr. Luca Delle  
Monache, and Dr. Sue Haupt!

---



Questions?