Basic information

1. Subject: Kansas (KS), winter wheat

2. Time: 1974 - 1980

3. Factors: Soil moisture, temperature, stalk number (stem-to-head ratio)

4. Source: USDA, US National Weather Service (over 400 stations)

Overall results

1. Y = a + bX: X is the number of stems, Y is the number of heads;

W = c: c is the average head weight

2. Role of soil moisture

- Initialization: data at the end of April

- SW = SW' + P - ET - D - R: soil moisture at SW current, soil moisture before SW', the difference of precipitation, evapotranspiration, subsurface drainage, surface runoff for P, ET, D and R (current - past)

- SW = SW' + P - ET - D: R can be ignored

- When SW' + P - ET < field capacity, D = 0; otherwise, SW = field capacity

- Relative soil water (RSW) = SW / field capacity ∈ [0,1]

3. ET = r(PET)

- Potential evapotranspiration (PET) = (α + βT)K, T is the mean maximum temperature and K is the large water surface free evaporation constant

- In this paper: K = 0.77

- r = f(RSW): Figure 5 (P17)

4. Field capacity: different depths at different times: 5 ft during hibernation (+5 ft storage), 10 ft in spring

5. RSW: negatively correlated with stem-to-head ratio (total number of heads still declining), positively correlated with head weight -> compensating effect

6. Temperature, number of stalks: same as RSW

7. Specific factors

- April relative soil moisture (RSWA)

- Cumulative maximum temperature (TEMP)

- Average stalk number (STKN)

- Mixed conditions (RSWA + TEMP ; STKN + TEMP ; RSWA + STKN)

Specific results

1. RSW(init) = 0.7: not sensitive, RSW converges within [0.5, 1]

2. Early season (April/May) soil moisture is important -> April/May soil moisture is predictive of outcome alone

3. Early season (April/May) temperature is not important -> cannot be stored, but can influence head number (energy distribution)

- Warmer - more stalks - stalk to head ratio close to 1

4. Remove 1980 anomaly -> better prediction of head counts

- Significant improvement for RSWA/RSWA+STKN

- 1977/1978/1979: greater improvement, 1980: better for time series

5. Remove higher data for 1976/1977 -> better prediction of 1980 head weights

- STKN/RSWA+STKN: better/as good as five-year time series (for all years)

6. RSWA does better in predicting head weight (even without excluding 1980)

7. Five year time series gives more conservative head weights, smoother results and uniformity of AAE

8. Use of TEMP and other indices not as effective in practice (head weight/head count)

9. WYS grouping for RSWA

- In this paper: 0 - 0.2, 0.2 - 0.6, 0.6 - 1

10. Significant improvement in head count and head weight by RSW grouping

- but no improvement in the 1980 anomaly year

Model analysis

1. Using: interannual stratification (BYS), intra-annual stratification (WYS); comparison with three-year time series

- Performs better when weather factors are important

- WYS is more expensive, requires more data -> in practice, difficult to outperform BYS

2. Consider the effect of weather on plant early characteristics and consider the effect of early plant characteristics on later periods

- Grouping of data by weather for different years

3. Model evaluation

- Average absolute error (AAE) = SUM(C'i - Ci) / n

- Does not disregard large errors in individual samples: simpler and more accurate

4. Specific algorithm

- For each year, choose the 3 years with the closest data

- Use these data to calculate a, b, c

- Calculate AAE: uncorrelated (weak relationship) -> randomly ranked, strong relationship -> centrally ranked (should reduce error)

- Can be calculated using a chain: early -> moisture/temperature - stalk ratio/stalk number, late -> stalk ratio/stalk number - head weight/head number (final harvest); stalk ratio positively correlated with stem number

5. WYS: split each year into multiple time groups

- For each time group, find the closest data to the weather conditions

6. Exclude anomalous data: less helpful for WYS - less sensitive

Cautions

1. 1980 anomaly to be removed: large difference between early and late season -> model results are inaccurate and affect other forecasts

2. KS: western semi-arid, midday sub-humid, eastern humid -> rich data

3. Time scale fits spatial scale: county level - one month

- Avoid over-fitting and useless calculations

4. Field capacity: the maximum amount of water available for plants that the soil can hold

5. Some other factors: disease, insects, anthropogenic intervention, etc. is not considered -> may be worse than 3 year time series

Summary

1. Chain analysis can be used to arrive at the final result, finding the most appropriate factors and weightings at each stage

2. For each condition, it is possible to consider it as a whole, separately, or only in part (to find the best combination)

3. Grouping before forecasting, splitting before modelling -> more accurate

4. Different distances, different parameters, different choices of points -> multiple predictions to be averaged (or weighted)

5. Further analysis of the influences of each factor - more accurate, more data

6. Remove anomalous data, learn multiple times, test different factors multiple times -> find the best model

7. Beware of false correlations in individual data (errors cancel each other out)

8. More complex models, more detailed considerations, not necessarily better -> consider the number of factors, amount of data that can be obtained, analytical power

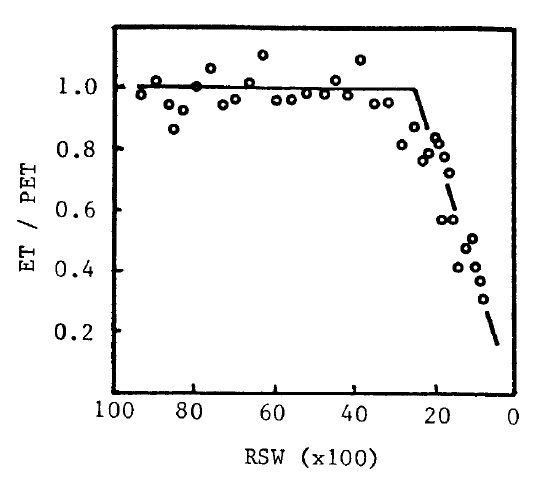


Figure 5. Ratio of evapotranspiration to potential evapotranspiration as influenced by relative soil water (Meyer and Green, 1980).