Spatial and Temporal-Spatial Analysis in R with applications to Crop yield analysis with climate data in Baden-Württemberg

Abstract

Spatial and Temporal-Spatial Analysis are statistical analysis tools for finding patterns, detect anomalies, or test hypotheses and theories based on spatial data. In our project, we use these tools for studying the relationship between crop yield such as Winterweizen and Wintergerste and climate factors among Baden-Wurttemberg during the time period from 2016 to 2020. The following parts of this report is organized as follows: firstly, background concepts and knowledge are put in the first paragraph. Spatial Models, concentrating on interpreting data and modeling data from a spatial perspective, is stated in the second paragraph. In final paragraph, we analyze the data from the perspective of both temporal and spatial. For each model we build in this report, we start with introducing definitions and concepts and followed by practical implementation in our project in R and model results and interpretation is stated as the concluding part.

Contents

Chapter 1 - Introduction

Knowledge Preparation

Spatial Data Analysis

Definition:

Spatial data analysis refers to a set of techniques designed to find pattern, detect anomalies, or test hypotheses and theories, based on spatial data. More rigorously, a technique of analysis is spatial if and only if its results are not invariant under relocation of the objects of analysis—in other words, that location matters [1]. Spatial data analysis is a combination of mathematics, geography, and computational science, which generally involves the process of collecting, creating information from a set of geographic features and evaluating, analyzing, or modeling the data in a systematic way in order to get insights for prediction and interpretation.

Spatial Autocorrelation

Definition:

The term spatial autocorrelation refers to the presence of systematic spatial variation in a mapped variable [2]. More precisely, spatial autocorrelation focuses on studying the dependence of space unit in the whole space and its neighboring space units on some particular variables aimed to analyze the properties of the spatial distribution of these space units [3]. Where adjacent observations have similar data values the map shows positive spatial autocorrelation, and the map shows negative spatial autocorrelation where adjacent observations tend to have very contrasting values. There are several statistical techniques for detecting its presence. The presence of spatial autocorrelation is important, (a) because it is usually taken as indicating that there is something of interest in the distribution of map values that calls for further investigation in order to understand the reasons behind the observed spatial variation, and (b) because the presence of spatial autocorrelation implies information redundancy and has important implications for the methodology of spatial data analysis [2].

Weight Matrix

The spatial weight matrix is a manifestation of the spatial structure of the data. It is a quantification of the spatial relationships that exist between data set elements (or, at least, a quantification of the method of conceptualizing such relationships). Because the spatial weight matrix imposes a structure on the data, you should choose a conceptualization that best reflects how the elements actually interact with each other (of course, you also need to consider what you are trying to measure). Although it is physically realized by various methods, conceptually, the spatial weight matrix is an NxN table (N represents the number of elements in the data set). Each element has only one row and one column. The pixel value of any given row/column combination is the weight, which can be used to quantify the spatial relationship between these row and column features [4].

At the most basic level, there are two strategies for quantifying the relationship between data elements by creating weights: binary or variable weights. For binary strategies (fixed distance, K nearest neighbor, Delaunay triangulation, adjacency, or space-time window), the feature is either a neighborhood (1) or not (0). For weighting strategies (inverse distance or indifferent areas), neighboring elements have different levels of effect (or influence), and weights are calculated to reflect this change [4].

Moran's I Value

In statistics, Moran's I is a measure of spatial autocorrelation developed by Patrick Alfred Pierce Moran.

Moran's I is defined as

$$I = \frac{N}{W} \frac{\sum_i \sum_j w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2}$$

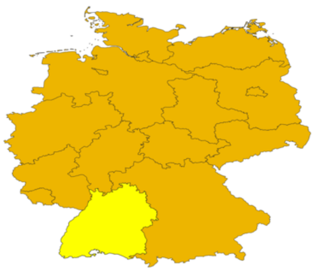
where N is the number of spatial units indexed by i and j ; x is the variable of interest; $\bar{x}$ is the mean of x ; w_{ij} is a matrix of spatial weights with zeroes on the diagonal (i.e., $w_{ii} = 0$); and W is the sum of all w_{ij}.

Software Introduction

QGIS

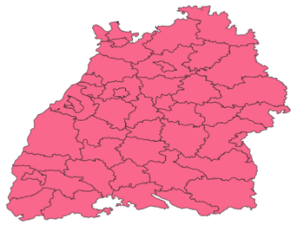
QGIS (until 2013 known as Quantum GIS) is a free and open-source cross-platform desktop geographic information system (GIS) application that supports viewing, editing, and analysis of geospatial data.

For example, we can have a view of the whole map of Germany and also a map with all divided cities.



Region of the whole country

Region of the whole country divided by cities



Region of the Baden-Württemberg divided by cities

In addition, it is also convenient to combined datasets in QGIS. For instances, we can join csv data containing variable information and shpfile containing geographical data together to make new shpfiles which is convenient to use in further analysis.

Chapter 2 - Spatial regression Models

Introduction

Regression methods are statistical tools to study relationship between one variable called dependent variable and other factors named independent variables. More specially, Spatial regression refers to Statistical method which, taking into consideration the location of each observation, measures the strength and direction of relationships between a dependent spatial variable and independent spatial variables. Because of the nature of combining spatial information with regression, the spatial regression can better explain the spatial relationship of the geographical phenomenon.

In this project, we do Spatial Regression Analysis for Winterweizen productions with the data from weather stations at the year of 2020, the same procedure applied for Wintergerste, which both of the code is stored in the file *Spatial Models.R*.

Since there are too many variable available for us, our model can only accommodate 6 variables, we do a variable selection procedure which figure out the top 6 significant variables we used in analysis.

Data Exploration

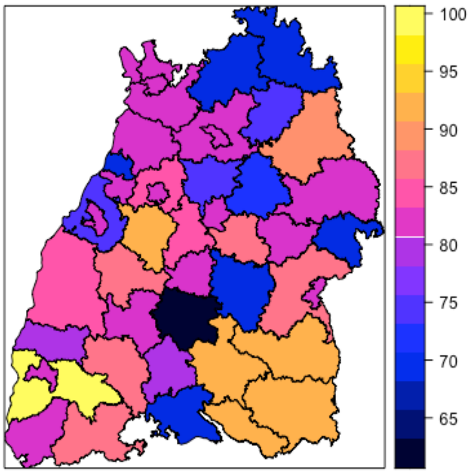
Firstly, import all the library we used here and below. We read the spatial data from *shpfile* using function *readOGR*. After that, we can exploration all variable names from the command *names(data)* and obtain statistical information from *summary(data)*. Visualization of the production data on map can be made by calling *spplot*.

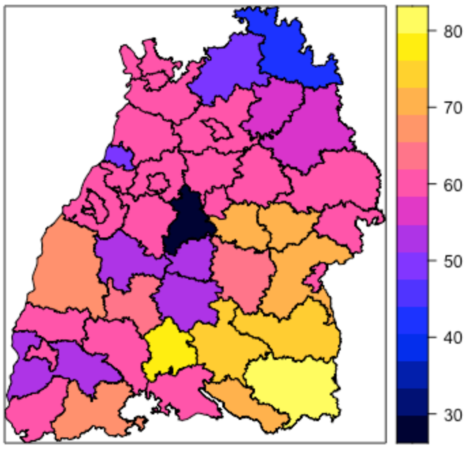
```
library(rgdal)
library(spdep)
library(rgeos)

#Reading in Shapefile with Data
data = readOGR(dsn="/Users/jiangxiaoyu/Desktop/Files for Spatial Analysis Project/Baden-Württemberg shpfile/Baden_Wurttemberg_Final.shp",layer="Baden_Wurttemberg_Final")
names(data) #X2020_csv1 to X2020_csv12 are: TEMP TEMP_ATTRIBUTES    DEWP DEWP_ATTRIBUTES STP WDSP    WDSP_ATTRIBUTES MXSPD   GUST    MAX MIN PRCP
            #X2020_csv13:Winterweizen
            #X2020_csv14:Wintergerste
summary(data)
spplot(data,"X2020_csv13") # make a map for Winterweizen
```

```
12. spplot(data,"X2020_csv14") # make a map for Wintergerste
```

The output picture is showed below:





Winterweizen yield in 2020 among different regions Wintergerste yield in 2020 among different regions

From Spatial plot made above, one information we get is that it seems that Wintergerste production is generally higher at the south-east part of Baden-Württemberg than other regions. However, generally speaking, we found the geographical pattern of Winterweizen and Wintergerste yield is not very clear, which means we may need more statistical tools to discover more inside.

Basic Ordinary Least Regression Model (OLS)

Definition

In statistics, ordinary least squares (OLS) is a type of linear least squares method for estimating the unknown parameters in a linear regression model. OLS chooses the parameters of a linear function of a set of explanatory variables by the principle of least squares: minimizing the sum of the squares of the differences between the observed dependent variable (values of the variable being observed) in the given dataset and those predicted by the linear function of the independent variable [5].

Geometrically, this is seen as the sum of the squared distances, parallel to the axis of the dependent variable, between each data point in the set and the corresponding point on the regression surface — the smaller the differences, the better the model fits the data. The resulting estimator can be expressed by a simple formula, especially in the case of a simple linear regression, in which there is a single regressor on the right side of the regression equation [5].

The OLS estimator is consistent when the regressors are exogenous, and—by the Gauss–Markov theorem—optimal in the class of linear unbiased estimators when the errors are homoscedastic and serially uncorrelated. Under these conditions, the method of OLS provides minimum-variance mean-unbiased estimation when the errors have finite variances. Under the additional assumption that the errors are normally distributed, OLS is the maximum likelihood estimator [5].

Formula for OLS model:

$$y = X\beta + \varepsilon$$

Same as

$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_m x_{im} + \varepsilon_i$ for $m = 1 \ldots M;\ i = 1 \ldots N$

Where N is the total number of observations and M is the number of features.

- y: dependent variable, in our project this term refers to Winterweizen or Wintergerste production.
- X: independent variables, contains the information of climate factors in our example, such as TEMP, WDSP etc.
- β: regression parameters, which represent the weight of each factor in X for explaining dependent variable y.
- ε: Normal Stochastic error term

Implementation

The first step we do is to select top 6 variables we will used for our Spatial Models. These are selected based on significance level in the Global Regression. Then, Spatial Information is stored in *listw1* which contained the Geographical information used for following steps. Ordinary Least Square model is established following the command *lm()*, whose descriptive data can be called by *summary()*. *lm.morantest* operates a statistical hypothesis test for testing spatial significance, where null hypothesis is assigned to be ‘no spatial relationship’.

```
#Creating Spatial Weights in R
```

```
queen.nb = poly2nb(data)
rook.nb = poly2nb(data,queen = FALSE)

# Regression for X2020_csv13(Winterweizen)
# Top 6 significant variable from Global regression: X2020_csv10, X2020_csv11, X2020_csv_2, X2020_csv12, X2020_csv_1, X2020_csv_7
# Bug here: we can only apply 6 independent variables here
```

```
reg.equ1 = X2020_csv13 ~ X2020_csv10 + X2020_csv11 +  X2020_csv_2 + X2020_cs
v12 + X2020_csv_1 + X2020_csv_7

#Convert queen.nb to queen.listw (needed for regression)
queen.listw = nb2listw(queen.nb)
rool.listw = nb2listw(rook.nb)
listw1 = queen.listw

#OLS for X2020_csv13 (Winterweizen)
reg1 = lm(reg.equ1,data=data)
summary(reg1)
lm.morantest(reg1,listw1) #null hypothesis: no spatial relationship
```

The p-value returned from *lm.morantest* is 0.3958, which is not significant even at 10% level. This fact may suggest that spatial information in data is not so clear in interpreting the regression.

Spatial Lag X (SLX)

The Spatial Lag Model contains the assumption that the values of dependent variable are influenced by other local independent variables. In our project, the yield of crops in one state is not only influence by the weather conditions in the state, but also that of nearby regions.

Formula: $y = X\beta + WX\theta + \epsilon$

- y: dependent variable, in our project this term refers to Winterweizen or Wintergerste production.
- X: independent variables, contains the information of climate factors in our example, such as TEMP, WDSP etc.
- β: regression parameters, which represent the weight of each factor in X for explaining dependent variable y.
- ε: Normal Stochastic error term
- θ: lag x parameters, which represent the weight of each factor in nearby regions for explaining dependent variable y.
- W: Spatial weight matrix which represents the neighborhood structure.

Implementation

```
#SLX Spatially Lagged X
```

```
reg2 = lmSLX(reg.equ1,data = data,queen.listw)
summary(reg2)
```

We call the command lmSLX to implement this model, the p-value from F-Statistics is 0.1096, and Adjusted R-squared is 0.1679.

Spatial Lag Autoregression (SAR)

The Spatial Lag Autoregression assumes that the values of dependent variable are connected with other local dependent variables. In our project, it is that the yield of crops in one state is related to that of nearby regions.

Formula: $y = X\beta + \rho W y + \epsilon$

- y: independent variable, in our project this term refers to Winterweizen or Wintergerste production.
- X: independent variables, contains the information of climate factors in our example, such as TEMP, WDSP etc.
- β: regression parameters, which represent the weight of each factor in X for explaining dependent variable y.
- ϵ: Normal Stochastic error term.
- ρ: the parameter of the spatially lagged dependent variable.
- W: Spatial weight matrix which represents the neighborhood structure.

Implementation

```
#SAR Spatial Lag Autoregression
reg3 = lagsarlm(reg.equ1,data = data,queen.listw)
summary(reg3)
```

We call the command *lagsarlm* to implement this model, LM test for residual autocorrelation returns test value: 2.7765 and p-value: 0.095659.

Spatial Error Model (SEM)

The Spatial Error Model reflects the neighboring error terms that may affect the error term in the region. It is equivalent to making the assumption that the residual value is a function not only of the unexplained random error term but also a function of the region's neighbors' residual values [3].

Formula: $y = X\beta + u, u = \lambda W u + \varepsilon, where\ \varepsilon \sim i.i.d.$

- y: independent variable, in our project this term refers to Winterweizen or Wintergerste production.
- X: independent variables, contains the information of climate factors in our example, such as TEMP, WDSP etc.
- β: regression parameters, which represent the weight of each factor in X for explaining dependent variable y.
- u: a function depends on both residual in the region and that of neighboring regions.
- λ: the parameter which represents the weight of each nearby stochastic error for explaining dependent variable u.
- W: Spatial weight matrix which represents the neighborhood structure.
- ε: Normal Stochastic error term.

Implementation

```
#SEM Spatial Error Model
reg4 = errorsarlm(reg.equ2,data=data,queen.listw)
summary(reg4)
```

Calling *errorsarlm* command will return the model results as LR test value: 1.3693, p-value: 0.24194.

Chapter 3 - Temporal-Spatial Models

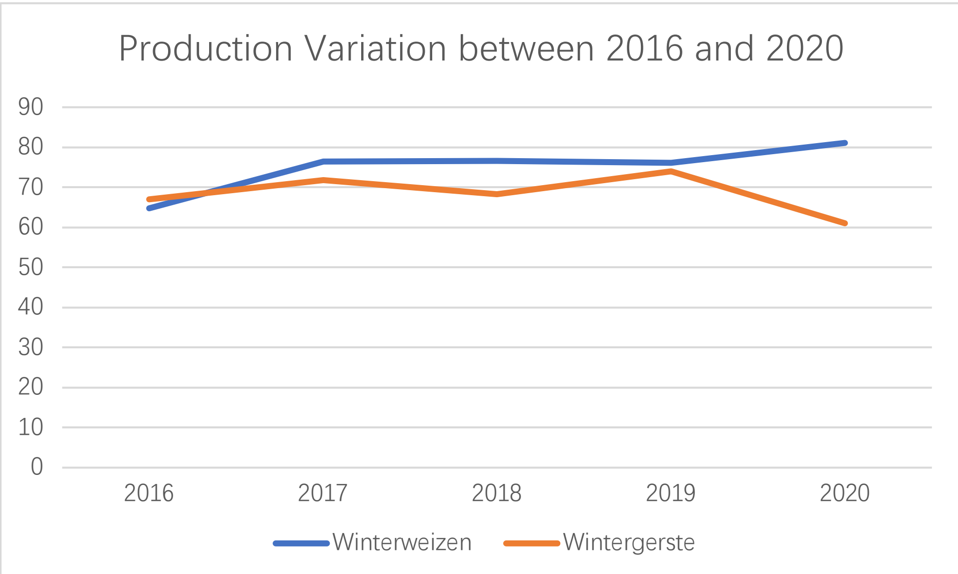
Introduction

Spatial refers to space. Temporal refers to time. Spatiotemporal, or spatial temporal, is used in data analysis when data is collected across both space and time. It describes a phenomenon in a certain location and time. Spatio-temporal data analysis is becoming a growing area of research with the development of powerful computing resources. Applications for spatio-temporal data analysis include the study of biology, ecology, meteorology, medicine, transportation and forestry.

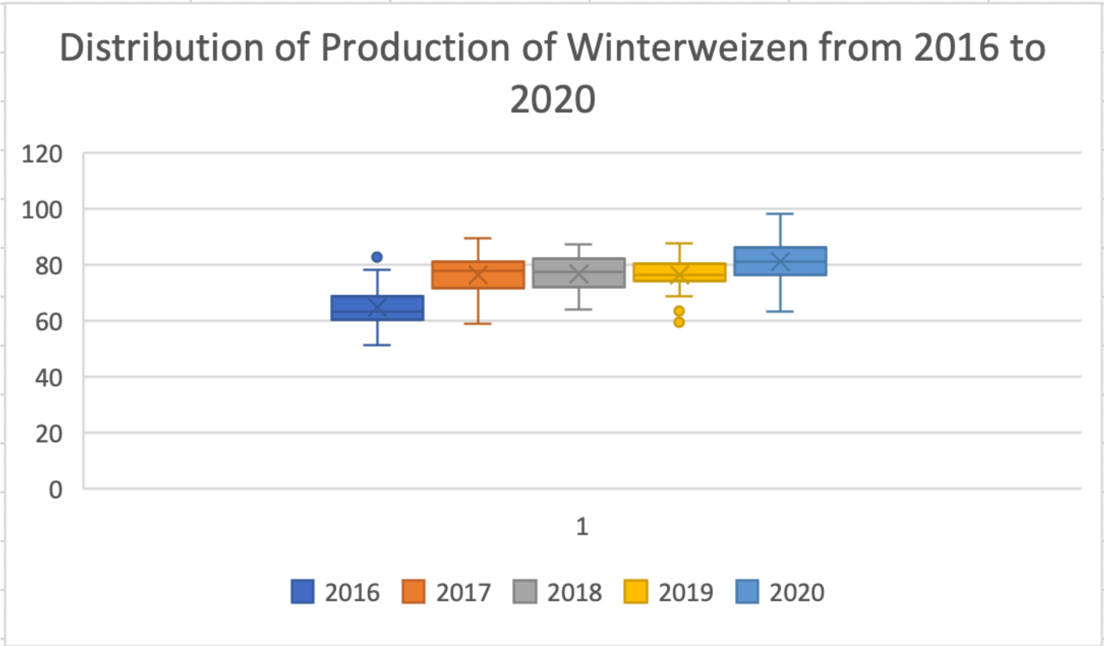
Temporal-Spatial Models is a class of statistical models which used the combined information from Space and Time to get insight about phenomenon we studied for. In our project, we do Temporal-Spatial analysis for Winterweizen and Wintergerste among the area of Baden-Württemberg and time period from 2016 to 2020, aiming to get useful finding from analysis.

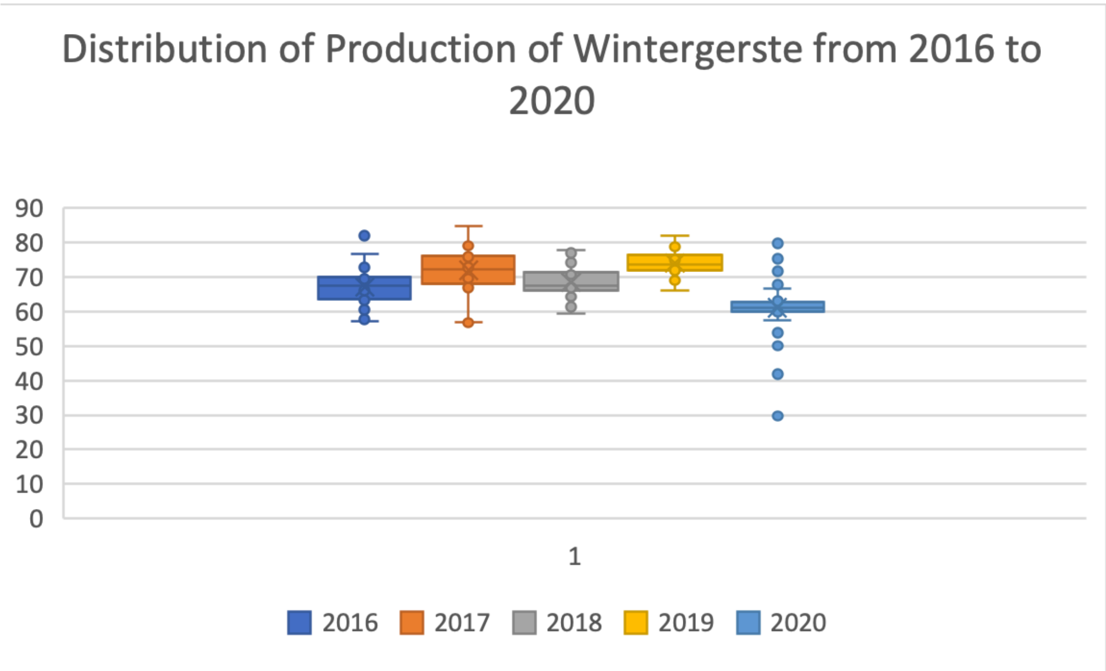
Data Exploration

In this section, we explore the yearly changed data by line chart and box plot. The line chart clearly shows how production of two crops varies among the five years. Generally, production for Winterweizen is increasing steadily, from around in 68 at 2016 to over 80 at 2020. In contrast, Wintergerste has a unstable change of production which is fluctuates between 60 and 75.



We also do box plot for these two crops each year. Overall, it seems that the distribution at the year of 2020 is more spread for both two crops.





Geographical and Temporal Weighted Regression (GTWR)

Model Formulation

$$y_i = \beta_0(u_i, v_i, t_i) + \sum_{k=1}^{m} \beta_k(u_i, v_i, t_i) x_{ik} + \varepsilon_i$$

- y_i: dependent variable, which is to be interpreted.
- (u_i, v_i, t_i): longitude, latitude and time index for i-th sample.

- β_k: regression coefficient for k-th explanatory variable and i-th sample, which depends on the value of (u_i, v_i, t_i).
- β_0: intercept coefficient.
- x_{ik}: the value taken at k-th explanatory variable and i-th sample.
- ε_i: Stochastic error term.

Variable Selection

In principal, we have 12 variables which can be used in the model. Different variable combinations may lead to different model robustness. Therefore, we are trying to search out a suitable set of variables which gives best interpretation for our data. Our approach is to order all 12 variables according to their significance in Global Regression model, and then fit 12 models each contains different number of variables (range from 1 to 12). After that, we collect all the summary data for each model, and compare models from different criterions. Finally, we decide the model with first 5 significant variables: MAX, MIN, TEMP_ATTRIBUTES, PRCP, TEMP is most suitable, since it has both smallest AIC value and largest Adjusted R-Square value.

Implementation

Here, we implement the model with the optimal set of variables described above. We do the analysis for Winterweizen at the first stage, and followed by analysis for Wintergerste.

The first step we do is to import all libraries we need for future use. Data are read and process to suitable data-type which is ready for being used in *gtwr* function from *GWmodel*. Then, optimal bandwidth is calculated from bw.gtwr command, which prepares for calling gtwr to fit in the model. Finally, the summary can be returned by calling its name. The same procedure can be applied for the analysis of Wintergerste.

```
library(GWmodel)
library(rgdal)

data_csv = read.csv("/Users/jiangxiaoyu/Desktop/21 Summer/Spatial_Regression/2016_2020data.csv", header=T)
names(data_csv)
coordinates(data_csv) <- ~LATITUDE+LONGITUDE

# Regression for Winterweizen
```

```
fixedgaussian_bw <- bw.gtwr(Winterweizen ~ TEMP + TEMP_ATTRIBUTES + DEWP + DEWP_ATTRIBUTES + STP + WDSP + WDSP_ATTRIBUTES, data=data_csv,
                            obs.tv = data_csv$YEAR, approach = "AICc", kernel = "gaussian", adaptive = F, lamda = 0.05, t.units = "YEAR", ksi = 0,
                            verbose = TRUE)

#Run the GTWR model with the fixed Gaussian bandwidth
gtwr_model <- gtwr(Winterweizen ~ TEMP + TEMP_ATTRIBUTES + DEWP + DEWP_ATTRIBUTES + STP + WDSP + WDSP_ATTRIBUTES, data=data_csv,
                   obs.tv = data_csv$YEAR, st.bw = fixedgaussian_bw, kernel="gaussian",
                   adaptive=FALSE, p=2, theta=0, longlat=F,lamda=0.05,t.units = "YEAR",ksi=0)

gtwr_model ## summary of model

#-----------------------------------------------------------------------------------------------------------------------------
# Regression for Wintergerste

fixedgaussian_bw2 <- bw.gtwr(Wintergerste ~ TEMP + TEMP_ATTRIBUTES + DEWP + DEWP_ATTRIBUTES + STP + WDSP + WDSP_ATTRIBUTES + MXSPD + GUST + MAX + MIN + PRCP, data=data_csv,
                            obs.tv = data_csv$YEAR, approach = "AICc", kernel = "gaussian", adaptive = F, lamda = 0.05, t.units = "YEAR", ksi = 0,
                            verbose = TRUE)

#Run the GTWR model with the fixed Gaussian bandwidth
gtwr_model2 <- gtwr(Wintergerste ~ TEMP + TEMP_ATTRIBUTES + DEWP + DEWP_ATTRIBUTES + STP + WDSP + WDSP_ATTRIBUTES + MXSPD + GUST + MAX + MIN + PRCP, data=data_csv,
                   obs.tv = data_csv$YEAR, st.bw = fixedgaussian_bw2, kernel="gaussian",
                   adaptive=FALSE, p=2, theta=0, longlat=F,lamda=0.05,t.units = "YEAR",ksi=0)

gtwr_model2 ## summary of model
```

Model Output (Result)

```
***********************************************************************
   *                       Package   GWmodel                          *
```

```
******************************************************************************
Program starts at: 2021-10-11 08:12:55
Call:
gtwr(formula = Winterweizen ~ TEMP + TEMP_ATTRIBUTES + DEWP +
 DEWP_ATTRIBUTES + STP + WDSP + WDSP_ATTRIBUTES, data = data_csv,
 obs.tv = data_csv$YEAR, st.bw = fixedgaussian_bw, kernel = "gaussian",
 adaptive = FALSE, p = 2, theta = 0, longlat = F, lamda = 0.05,
 t.units = "YEAR", ksi = 0)

Dependent (y) variable:  Winterweizen
Independent variables:  TEMP TEMP_ATTRIBUTES DEWP DEWP_ATTRIBUTES STP WDSP WDSP_ATTRIBUTES
Number of data points: 220
******************************************************************************
*                     Results of Global Regression                           *
******************************************************************************

Call:
 lm(formula = formula, data = data)

Residuals:
   Min       1Q   Median       3Q      Max
-21.6000  -5.9319   0.6199   5.8526  21.9488

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)     -112.7217   183.2332  -0.615  0.53909
TEMP               1.1382     0.4151   2.742  0.00663 **
TEMP_ATTRIBUTES    8.3650   211.3414   0.040  0.96846
DEWP              -0.2678     0.1301  -2.059  0.04067 *
DEWP_ATTRIBUTES  -74.3338    58.3982  -1.273  0.20446
STP                0.1230     0.1781   0.691  0.49034
WDSP               1.0423     0.7302   1.427  0.15495
WDSP_ATTRIBUTES   66.4899   202.0800   0.329  0.74246

---Significance stars
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 8.247 on 212 degrees of freedom
Multiple R-squared: 0.1212
Adjusted R-squared: 0.09221
F-statistic: 4.178 on 7 and 212 DF,  p-value: 0.0002476
***Extra Diagnostic information
Residual sum of squares: 14419.54
```

```
   Sigma(hat): 8.132939
   AIC:  1562.53
   AICc:  1563.387
   ***********************************************************************
   *    Results of Geographically and Temporally Weighted Regression     *
   ***********************************************************************

   ********************Model calibration information*********************
   Kernel function for geographically and temporally weighting: gaussian
   Fixed bandwidth for geographically and temporally weighting:  0.5942687
   Regression points: the same locations as observations are used.
   Distance metric for geographically and temporally  weighting: A distance matrix is specified for this model calibration.

   ****************Summary of GTWR coefficient estimates:*****************
                            Min.     1st Qu.      Median     3rd Qu.      Max.
   Intercept       -6.8880e+02 -3.5976e+02  1.7447e+02  2.3864e+02  342.3691
   TEMP            -1.9986e+00 -1.1623e+00 -3.0909e-01  5.0266e-01    5.9910
   TEMP_ATTRIBUTES -7.6816e+02 -3.5693e+02  4.0343e+02  6.6985e+02 1208.5473
   DEWP            -1.5864e+00 -2.8116e-01  6.4435e-02  4.1051e-01    2.0011
   DEWP_ATTRIBUTES -7.0413e+02 -5.8171e+01  8.6772e+00  1.7529e+02  515.2550
   STP             -2.6204e-01 -1.0160e-01 -5.0538e-02  3.2754e-01    0.5090
   WDSP            -1.7299e+00 -6.6336e-01  2.8280e-01  3.7156e+00    5.4117
   WDSP_ATTRIBUTES -1.3864e+03 -6.4474e+02 -4.3872e+02  3.7287e+02  850.4348
   ***********************Diagnostic information*************************
   Number of data points: 220
   Effective number of parameters (2trace(S) - trace(S'S)): 31.17374
   Effective degrees of freedom (n-2trace(S) + trace(S'S)): 188.8263
   AICc (GWR book, Fotheringham, et al. 2002, p. 61, eq 2.33): 1495.542
   AIC (GWR book, Fotheringham, et al. 2002,GWR p. 96, eq. 4.22): 1457.884
   Residual sum of squares: 8598.446
   R-square value:  0.475984
   Adjusted R-square value:  0.3890124
```

```

   ***********************************************************************
```

```
***********************************************************************
   *                       Package   GWmodel                          *
   ***********************************************************************
   Program starts at: 2021-10-11 08:20:03
   Call:
   gtwr(formula = Wintergerste ~ TEMP + TEMP_ATTRIBUTES + DEWP +
    DEWP_ATTRIBUTES + STP + WDSP + WDSP_ATTRIBUTES + MXSPD +
    GUST + MAX + MIN + PRCP, data = data_csv, obs.tv = data_csv$YEAR,
    st.bw = fixedgaussian_bw2, kernel = "gaussian", adaptive = FALSE,
    p = 2, theta = 0, longlat = F, lamda = 0.05, t.units = "YEAR",
    ksi = 0)

   Dependent (y) variable:  Wintergerste
   Independent variables:  TEMP TEMP_ATTRIBUTES DEWP DEWP_ATTRIBUTES STP WDSP WDSP_ATTRIBUTES MXSPD GUST MAX MIN PRCP
   Number of data points: 220
   ***********************************************************************
   *                    Results of Global Regression                     *
   ***********************************************************************

   Call:
    lm(formula = formula, data = data)

   Residuals:
    Min       1Q   Median      3Q     Max
-37.914   -3.729    0.317   4.137  14.873

   Coefficients:
                     Estimate Std. Error t value Pr(>|t|)
   (Intercept)      473.25733  204.95265   2.309 0.021924 *
   TEMP              -1.53635    0.41132  -3.735 0.000243 ***
   TEMP_ATTRIBUTES -382.69887  206.23648  -1.856 0.064928 .
   DEWP              -0.16077    0.11480  -1.400 0.162875
   DEWP_ATTRIBUTES  -67.14460   51.30118  -1.309 0.192043
   STP               -0.25137    0.19805  -1.269 0.205789
   WDSP              -0.94747    0.87445  -1.083 0.279849
   WDSP_ATTRIBUTES  448.15075  203.39074   2.203 0.028670 *
   MXSPD              0.20110    0.46908   0.429 0.668572
   GUST              -0.03493    0.01382  -2.527 0.012237 *
   MAX                0.07111    0.12000   0.593 0.554106
```

```
MIN                -0.11617    0.11501   -1.010 0.313631
PRCP                0.08479    0.05164    1.642 0.102160

---Significance stars
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 6.947 on 207 degrees of freedom
Multiple R-squared: 0.1351
Adjusted R-squared: 0.08497
F-statistic: 2.695 on 12 and 207 DF,  p-value: 0.002143
***Extra Diagnostic information
Residual sum of squares: 9989.413
Sigma(hat): 6.769268
AIC:  1491.777
AICc:  1493.826
***********************************************************************
*    Results of Geographically and Temporally Weighted Regression     *
***********************************************************************

*********************Model calibration information*********************
Kernel function for geographically and temporally weighting: gaussian
Fixed bandwidth for geographically and temporally weighting:  0.6824789
Regression points: the same locations as observations are used.
Distance metric for geographically and temporally  weighting: A distance matrix is specified for this model calibration.

****************Summary of GTWR coefficient estimates:*****************
                           Min.     1st Qu.      Median     3rd Qu.      Max.
Intercept       -7.4152e+01  5.0540e+02  1.0497e+03  1.6547e+03 2847.1360
TEMP            -1.0835e+01 -6.4133e+00 -3.0595e+00 -1.5332e+00   -0.7078
TEMP_ATTRIBUTES -9.8727e+02 -3.4611e+02 -1.4543e+02  3.6082e+01 1690.1246
DEWP            -4.6403e-01 -6.9113e-02  5.8565e-01  5.2869e+00   11.2826
DEWP_ATTRIBUTES -2.1550e+02 -1.2134e+02 -4.3988e+01  5.4560e+01  296.6603
STP             -2.4886e+00 -1.3435e+00 -6.5564e-01 -3.5116e-01    0.2100
WDSP            -7.9485e+00 -4.2342e+00 -1.3646e+00  1.1632e+00    4.1475
```

```
    WDSP_ATTRIBUTES -1.5505e+03 -5.1027e+01  1.9979e+02  2.9570e+02  979.7042
    MXSPD            -3.7167e+00  4.8705e-02  5.5875e-01  2.5077e+00   10.4988
    GUST             -1.2147e-01 -9.8468e-02 -5.8121e-02 -2.9943e-02    0.0130
    MAX              -1.8294e-01 -2.2804e-02  7.8308e-01  2.4241e+00    4.5332
    MIN              -5.0296e+00 -2.4264e+00 -9.0139e-01 -4.3425e-02    0.0547
    PRCP              1.3273e-02  5.2020e-02  1.5239e-01  2.6474e-01    0.3968
    ************************Diagnostic information*************************
    Number of data points: 220
    Effective number of parameters (2trace(S) - trace(S'S)): 39.07315
    Effective degrees of freedom (n-2trace(S) + trace(S'S)): 180.9269
    AICc (GWR book, Fotheringham, et al. 2002, p. 61, eq 2.33): 1429.024
    AIC (GWR book, Fotheringham, et al. 2002,GWR p. 96, eq. 4.22): 1377.733

    Residual sum of squares: 5765.95
    R-square value:  0.5007814
    Adjusted R-square value:  0.3923704

    ***********************************************************************
```

The GTWR model was implemented for Winterweizen and Wintergerste data, the Adjusted R-square value are 0.3890124 and 0.3923704 respectively, which is significantly larger compared to 0.09221 of global regression model (without spatial information), showing that geographical data is useful here.

Reference Lists:

1. Goodchild M. (2008) Data Analysis, Spatial. In: Shekhar S., Xiong H. (eds) Encyclopedia of GIS. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-35973-1_236
2. R.P. Haining, in International Encyclopedia of the Social & Behavioral Sciences, 2001 https://www.sciencedirect.com/topics/computer-science/spatial-autocorrelation
3. Jingyi Hu (2019) Spatial Analysis in R with the application about shrimp farms risk assessment in Indonesia
4. ArcGis https://pro.arcgis.com/zh-cn/pro-app/latest/tool-reference/spatial-statistics/how-generate-spatial-weights-matrix-spatial-statis.htm
5. Wikipedia https://en.wikipedia.org/wiki/Ordinary_least_squares

Appendix:

Variable Explanations

TEMP: Temperature
TEMP_ATTRIBUTES: Temperature(transformed)
DEWP: dew point
DEWP_ATTRIBUTES: dew point(transformed)
STP: station pressure
WDSP: wind direction speed
WDSP_ATTRIBUTES: wind direction speed (transformed)
MXSPD: max wind speed
GUST: gust
MAX: Max Temperature
MIN: Min Temperature
PRCP: precipitation
Winterweizen: production of Winterweizen
Wintergerste: production of Wintergerste