Available online at www.isas.org.in/jisas
JOURNAL OF THE INDIAN SOCIETY OF
AGRICULTURAL STATISTICS 72(2) 2018 165–174

Knowledge Engineering for Apportioning District Level Data in Agriculture

Saravanakumar R[1], Rajni Jain[2], Alka Arora[1] and Sudeep Marwaha[1]
[1] ICAR-Indian Agricultural Statistical Research Institute, New Delhi
[2] ICAR-National Institute for Agricultural Econimcs and Policy Research, New Delhi

Received 17 November 2017; Revised 10 July 2018; Accepted 11 July 2018

SUMMARY

Agriculture is the backbone of India and data is a driver of growth and change. In agriculture domain also, large amount of data is being collected by various agencies including Government of India. The data is generated every day at various levels namely household, village, district, state and country. Data collection is done and recorded at current status regarding spatial boundaries of the units. At country level, there is no change in the boundary. State level spatial changes are rare and easy to capture in the data analysis. But, district level temporal data analysis creates problem because of the variation in number of districts and changes in their areas over a period of time. Thus, temporal data at district level requires adjustment called apportioning before analysis. These adjustment in data vary from variable to variable. The paper discusses apportioning methodology for different categories of variables using mathematical notations as well as using software named DAS developed for this purpose. The software has been developed using the programming language C# and tested using different datasets. The software called DAS (http://das.iasri.res.in) has been developed and also made available online.

Keywords: Apportioning, DAS, District boundaries, Web based and district data.

1. INTRODUCTION

Data is a driver of growth and change. Data analysis helps to take right decision regarding creation of new infrastructure, new businesses, new monopolies, and also for agriculture and rural farm welfare. However, analysis has to be done on the basis of appropriate administrative unit.States appears to be appropriate administrative unit to study regional variations in many aspects. However, agriculture performance generally differs widely within state due to varying regional characteristics in terms of resource endowments and climate. Therefore, need for lower administrative unit becomes apparent. Recognizing the importance of district level approach for agriculture development, NITI Aayog (earlier called Planning Commission) has asked the states to prepare district level plan for agriculture to get funding for development of agriculture sector during XI plan (Chand *et al.*, 2009). The main problem in temporal analysis at district level is- (i) number of districts in a state in time series data are different, (ii) geographical area belonging to a particular district is typically not constant over time. During a time period covered by any temporal data set, there may be numerous adjustments in the boundaries of some of the districts. The paper presents knowledge engineering for apportioning the data based on dynamics of district boundaries. Software DAS (District Data Apportioning Software) has been developed for apportioning of different types of variables. Knowledge engineering for this software has been presented with the help of mathematical notations also.

A time series is a sequence of numerical data in which each item is associated with a particular instant in time. Time series analysis requires handling of variations in geographical units boundaries over a period of time. If the boundaries are not constant, data analysis may show absurd trends. For example,

Corresponding author: Rajni Jain
E-mail address: rajnijain67@gmail.com

suppose population trends are under analysis for a district A. Population data analysis revealed that population of this district has decreased in 2001 census as compared to previous census which is reverse trend when compared with other neighbouring districts. Further probing highlighted that district area has decreased because part of this district has been made a separate new district 'B'. Such kind of problems are quite common in temporal analysis at district level. The problems becomes manifold, if analysis is to be done for all the districts in the country since independence. Different researchers have been tackling the problem in different ways as no standard solution is available. Some of these attempts are discussed in review section. This paper attempts to contribute by providing (i) knowledge engineering for this problem of data apportioning , so that researchers can apply it in their own language or software (ii) an online software by which researchers can get their data apportioned based on the base year, they want to analyse.

1.1 Review

A few researchers have worked on temporal district level analysis and apportioned the data manually using spread sheet. In 1975 a joint Jawaharlal Nehru University-Planning Commission project was initiated to study at the disaggregated district level, the changes that had taken place in agriculture in India during the period 1962-65 to 1970-73 as a consequence of the introduction of new-seed fertilizer technology during the mid-sixties. The report of the study was subsequently published (Bhalla and Alagh 1979, Bhalla and Tyagi 1989, Bhalla and Singh 2001, Bhalla and Singh 2010). However, apportioning methodology is not available in literature.

Village dynamics in South Asia knowledge bank (VDSA Project Report 2012 available at http://www.icrisat.org/vdsa/Include/Document/all-apportioned-web-document.pdf) has been developed by ICRISAT-ICAR-IRRI Collaborative Research Project. The database is maintained using two sets of district level time series data called apportioned database (based on 1966 districts) and un-apportioned database. In the apportioned database, data for all the variables for the districts formed after 1966 have been returned to their parent district and newly formed districts have been removed from the database (http://vdsa.icrisat.ac.in/). Kumar and Jain, 2013 examined the trends in growth and instability in Indian agriculture at the district level and also identified distinctive features and drivers of productivity growth across districts. The study used time series data from 1989 - 2009. The authors apportioned district data on the basis year of 1989. The methodology in the form of stand-alone application or in web application is not available in literature or on Internet till date.

The Indian states have been the standard unit of analysis for research on India that uses official data sources. Further, the boundaries of many states have been unchanged for over half a century and those of all major states were largely unchanged between 1971 and 2000. This stability has resulted in the relatively easy construction and use of panel data sets at the state level. Most district-level studies however have relied on cross-sectional analysis because district comparisons over time are complicated by multiple boundary changes (Kumar and Rohini, 2009). Between 1971 and 2001, the number of districts increased from 356 to 593, a rise of about 67 percent.Unchanged districts are single parent and single child districts and partitioned districts which share a single parent. For these two classes of districts it is straightforward to construct a district panel based on district boundaries corresponding to 1971, 1981, 1991, 2001 as a base year. They used manually calculated possible method of apportioning on 10 years basis. However, most of the agricultural data is compiled on annual basis which further increases the complexity in analysing district level time series data.

Apportioning is the most important process to assign and divide according the rule of proportion. The process of apportioning is tedious and time consuming especially when the time period under consideration is large along with large number of districts. Further, the problem is aggravated when numbers of variables to be apportioned are more. The problem is computation intensive and similar calculations are required to be repeated for each district every year with consideration of further possible changes in their boundaries. Manual computation using calculator or spreadsheet is prone to error, thus providing the scope for development of an online application. Data apportioning functions or utilities are not available in any of the popular statistical packages like SPSS, SAS, STATA, R Software or any other software to the best of our knowledge.

Jain *et al.* 2013 developed software for computing Total Factor productivity (WBSTFP) at any spatial levelincluding district level using time series data. WBSTFP softwaredemands pre-processed apportioned data and does not have facility of apportioning in the software. Apportioning refers to assign and divide according the rule of proportion. Manual computation using calculator or spreadsheet is prone to error, thus providing the scope for development of an online application. The proposed web based application for apportioning temporal data at district level is referred as DAS in the paper.

1.2 Proposed system

Based on the background above, it is anticipated that a data apportioning software which can support excel file spreadsheet data and provide option to apportion temporal district level data with reference to user specified base year will be helpful for policy makers, government and agricultural researchers in district level studies. To reduce the shortcomings of the manual method or by using spreadsheet computations, it is expected that the proposed software DAS should have the following features: (i) compatible with any operating system (ii) creation of user account (iii) online availability (iv) user validation (v) time saving (vi) capability to handle different type of data e.g absolute and density variables (Table 1) (vii) online help facility (viii) information on district boundary dynamics (ix) data validation (x) instantaneous online availability of output (xi) export option of apportioned output into Excel format for further analysis (xii) availability on web (xiii) user should be free from burden of download, install and upgrading.

The input data should be validated for compulsory attribute like district Id and year. Further, option for selection of categories of given attributes in the dataset to the two different categories namely absolute variable or density variable should be provided (Table 1). The proposed software has been developed using Visual Studio (Griffiths *et al.* 2003). The proposed web based application for apportioning temporal data at district level is refereed as DAS in the paper.

Table 1. Description of absolute and density type attributes

Categories of Attribute	Description	Examples
Absolute Category	The variables which requires adjustment in their values with changes in area and boundaries of districts	Area, population, production, animal population, irrigated area, total cultivated area, gross cropped area, net irrigated area, net cultivated area, gross irrigated area, waste land, degraded land, bovine population, total food grain production, pulse production, oil seed production, total sugar production, total cereal production, total vegetable production, total fruit production forest area, etc.
Density Category	The variables which are independent of area changes. The value of such a variable does not change with district boundaries	Yield, rainfall, temperature, population density, relative humidity etc.

1.3 Mathematical Foundation

Adjustment in the two categories of variables in time series data is required for various types of changes in the boundaries of districts as specified in Table 2.

Table 2. Different cases of apportioning method

S.no	Case ID	Description
1	Split n	A single district in the past has been divided into n districts
2	Merge n	Few districts (say n) in the past been have merged into one district
3	Merge-Split n	A new districts has been formed by merging adjacent area extracted from each of n districts
4	Multiple Split n	Few districts (say n) in the past has been merge and spilted into several districts (say m) in year C. And further merge and spilted into several more districts (say o) in year D

For each case described above, user may be interested in apportioning of data based on historical or recent boundary of districts.

1.3.1 Absolute category attribute

The attributes which require changesin their values as a consequence of changes in district boundary are referred as Absolute category attributes.

Let V_a is an absolute category variable and

$V_{ai}X$represents its value for i[th] year for any district X before apportioning. Let us further assume that $V`_{ai}X$represents its value for i[th] year for the district X after apportioning. It should be noted that for each type of split, depending on whether the base year is the historical year (A) or recent year, apportioning method will be different.

CASE 1 - SPLIT $_n$

(A) Historical year as the base year

Let district X splits into n districts say $X_1, X_2, X_3 \ldots X_n$ in a year C then:

$V`_{ai}X = V_{ai}X_p$(when i < C the year of change)

$V`_{ai}X = \sum_{p=1}^{n} V_{ai}X_p$ (when i ≥C the year of change)

X_p represents each of the splitted districts of the district X

(B) Recent year as the base year

In this case number of districts before the year of change will be increased. Thus each district X will be replaced by $X_1, X_2, X_3 \ldots X_n$ mathematically.

$V`_{ai}X_p = V_{ai}XR_p$ (when i < C the year of change)

$V`_{ai}X_p = V_{ai}X_p$ (when i ≥ C the year of change)

R_K represents each of the percentage of districts area contributions from district X

CASE 2 - MERGE $_n$

(A) Historical year as the reference year

Let district $X_1, X_2, X_3 \ldots X_n$ merge to form district X in year C then

$V`_{ai}X_p = V_{ai}X_p$ (when i < C the year of change)

$V`_{ai}X_p = V_{ai}XR_p$ (when i ≥ C the year of change)

(B) Recent year as the base year

In this case number of districts before the year of change will be decreased. Thus each and every district $X_1, X_2, X_3 \ldots X_n$ will be merged by X mathematically.

$V`_{ai}X = \sum_{p=1}^{n} V_{ai}X_p$ (when i < C the year of change)

$V`_{ai}X = V_{ai}X_p$(when i ≥C the year of change)

CASE 3 – MERGE-SPLIT $_n$

(A) Historical year as the base year

Let district $X_1, X_2, X_3 \ldots X_n$ merge and spilt to form $Y_1, Y_2, Y_3 \ldots Y_m$ in year C then

$V`_{ai}X_p = V_{ai}X_p$ (when i < C the year of change)

$V`_{ai}X_p = \sum_{q=1}^{m} V_{ai}Y_qR_q$ (when i ≥C the year of change)

(B) Recent year as the base year

Let district $X_1, X_2, X_3 \ldots X_n$ merge and spilt to form $Y_1, Y_2, Y_3 \ldots Y_m$ in year C then

$V`_{ai}Y = \sum_{p=1}^{n} V_{ai}X_pR_p$ (when i < C the year of change)

$V`_{ai}Y_q = V_{ai}Y_q$ (when i ≥C the year of change)

CASE 4 – MULTIPLE SPILT $_n$

(A) Historical year as the base year

Let district $X_1, X_2, X_3 \ldots X_n$ merge and spilt to form $Y_1, Y_2, Y_3 \ldots Y_m$ in year C and further it merge and split into $Z_1, Z_2, Z_3 \ldots Z_o$ in year D then

$V`_{ai}X_p = V_{ai}X_p$ (when i < C the year of change)

$V`_{ai}X_p = \sum_{q=1}^{m} V_{ai}Y_qR_q$ (when D>i ≥C the year of change)

$V`_{ai}X_p = \sum_{r=1}^{o} V_{ai}Z_oR_o$ (when i ≥ D the year of change)

(B) C to D year as the base year

Let district $X_1, X_2, X_3 \ldots X_n$ merge and spilt to form $Y_1, Y_2, Y_3 \ldots Y_m$ in year C and further it merge and split into $Z_1, Z_2, Z_3 \ldots Z_o$ in year D then

$V`_{ai}Y_q = \sum_{p=1}^{n} V_{ai}X_pR_p$ (when i < C the year of change)

$V`_{ai}Y_q = V_{ai}Y_q$ (when D>i ≥C the year of change)

$V`_{ai}Y_q = \sum_{r=1}^{o} V_{ai}Z_rR_r$ (when i ≥ D the year of change)

(C) Recent year as the base year

Let district $X_1, X_2, X_3 \ldots X_n$ merge and spilt to form $Y_1, Y_2, Y_3 \ldots Y_m$ in year C and further it merge and split into $Z_1, Z_2, Z_3 \ldots Z_o$ in year D then

$V`_{ai}Z_r = \sum_{p=1}^{n} V_{ai}X_pR_p$ (when i < C the year of change)

$V`_{ai}Z_r = \sum_{L=1}^{y} V_{ai} Y_q R_q$ (when D>i ≥C the year of change)

$V`_{ai}Z_r = V_{ai}Z_r$ (when i ≥ D the year of change)

1.3.2 Density category attribute

The attributes which do not change this values as a consequence of district boundary changes are referred as density category attributes.

Let V_a is a density category variable

$V_{ai}X$represents its value for i[th] year for any district X before apportioning

$V_a`_iX$represents its value for i[th] year for any district X after apportioning

CASE 1 - SPLIT $_n$

(A) Historical year as the base year

Let district X splits $X_1, X_2, X_3 \dots X_n$ in year C then

$V`_{ai}X = V_{ai}X_p$ (when i < C the year of change)

$V`_{ai}X = \sum_{p=1}^{n} V_{ai}X_p / n$ (when i ≥C the year of change)

(B) Recent year as the base year

In this case number of districts before the year of change will be increased. Thus each district X will be replaced by $X_1, X_2, X_3 \dots X_n$ mathematically.

$V`_{ai}X_p = V_{ai}X$ (when i < C the year of change)

$V`_{ai}X_p = V_{ai}X_p$ (when i ≥ C the year of change)

CASE 2 - MERGE $_n$

Historical year as the base year

Let district $X_1, X_2, X_3 \dots X_n$ merge to form district X in year C then

$V`_{ai}X_p = V_{ai}X_p$ (when i < C the year of change)

$V`_{ai}X_p = V_{ai}X$ (when i ≥ C the year of change)

(B) Recent year as the base year

In this case number of districts before the year of change will be decreased. Thus each and every district $X_1, X_2, X_3 \dots X_n$ will be merged by X mathematically.

$V`_{ai}X = \sum_{p=1}^{n} V_{ai}X_p / n$ (when i < C the year of change)

$V`_{ai}X = V_{ai}X_p$ (when i ≥C the year of change)

CASE 3 –MERGE-SPLIT $_n$

(A) Historical year as the base year

Let district $X_1, X_2, X_3 \dots X_n$ merge and spilt to form $Y_1, Y_2, Y_3 \dots Y_m$ in year C then

$V`_{ai}X_p = V`_{ai}X_p$ (when i < C the year of change)

$V`_{ai}X_p = \sum_{q=1}^{m} V_{ai}Y_q / m$ (when i ≥C the year of change)

(B) Recent year as the base year

Let district $X_1, X_2, X_3 \dots X_n$ merge and spilt to form $Y_1, Y_2, Y_3 \dots Y_m$ in year C then

$V`_{ai}Y = \sum_{p=1}^{n} V_{ai}X_p / n$ (when i < C the year of change)

$V`_{ai}Y_q = V`_{ai}Y_q$ (when i ≥C the year of change)

CASE 4 – MULTIPLE SPILT $_n$

(A) Historical year as the base year

Let district $X_1, X_2, X_3 \dots X_n$ merge and spilt to form $Y_1, Y_2, Y_3 \dots Y_m$ in year C and further it merge and split into $Z_1, Z_2, Z_3 \dots Z_o$ in year D then

$V`_{ai}X_p = V_{ai}X_p$ (when i < C the year of change)

$V`_{ai}X_p = \sum_{q=1}^{m} V_{ai}Y_q / m$ (when D>i ≥C the year of change)

$V`_{ai}X_p = \sum_{r=1}^{o} V_{ai}Z_r / o$ (when i ≥D the year of change)

(B) C to D year as the base year

Let district $X_1, X_2, X_3 \dots X_n$ merge and spilt to form $Y_1, Y_2, Y_3 \dots Y_m$ in year C and further it merge and split into $Z_1, Z_2, Z_3 \dots Z_o$ in year D then

$V`_{ai}Y_q = \sum_{p=1}^{n} V_{ai}X_p / n$ (when i < C the year of change)

$V`_{ai}Y_q = V_{ai}Y_q$ (when D>i ≥C the year of change)

$V`_{ai}Y_q = \sum_{r=1}^{o} V_{ai}Z_r / o$ (when i ≥ D the year of change)

(C) Recent year as the base year

Let district $X_1, X_2, X_3 \ldots X_n$ merge and spilt to form $Y_1, Y_2, Y_3 \ldots Y_m$ in year C and further it merge and split into $Z_1, Z_2, Z_3 \ldots Z_o$ in year D then

$V`_{ai}Z_r = \sum_{p=1}^{n} V_{ai}X_p / n$ (when i < C the year of change)

$V`_{ai}Z_r = \sum_{q=1}^{m} V_{ai}Y_q / m$ (when D>i ≥C the year of change)

$V`_{ai}Z_r = V_{ai}Z_r$ (when i ≥ D the year of change).

2. DAS PROCESS MODEL

A software process model is an abstract representation of the architecture, design or definition of the software process (Sommerville, 2011). There are varieties of software development process models to show how organizing the process activities can make the development more effective. One of the basic software process models is waterfall model. It is not flexible. Its phases are strictly linear (Munassar and Govardhan, 2010). So, the Royce's modified final waterfall model has been used in the development of the data apportioning software for district. The modified waterfall model uses the same phases as the pure waterfall but the phases are permitted to overlap in modified waterfall model.

2.1 Architecture and design of DAS

DAS is designed and implemented as a three layered structure with each layer corresponding to a different functionality (Grove, 2010). Three tier architecture includes Client Side Interface layer (CSIL), Server Side Application layer (SSAL) and Database layer. The CSIL is designed using Hyper Text Markup Language (HTML) (Powell, 2003), Cascading Style Sheet (CSS) (Bos and lie, 1999) and User Interface (UI).SSAL is designed using ASP.NET framework along with C# (Macdonald, 2002). Database transaction has been done using Structured Query Language and Microsoft SQL server 2012. Database contains tables like Ratio_ Table, User_Data and District used for apportioning data and one separate table for Login user. Login table facilitates tracking of users of the system. Ratio table stores every information about the District Id and the respective parent district id with ratio for each base year. The userdata table represents a temporary table to get data from dataset. District data table contain data about district id, state name, district name and its image (map of the districts). The entity-relationship data model describes the tables and their attributes involved in database and their relationships. The entity-relationship of the DAS software database is shown in Fig. 1.

2.2.1 Module Design

Various modules developed for the software are presented in Table 3. The module name represents the functionality managed by the corresponding module. For example, the apportioning module implements the apportioning of the district data using ratio table.

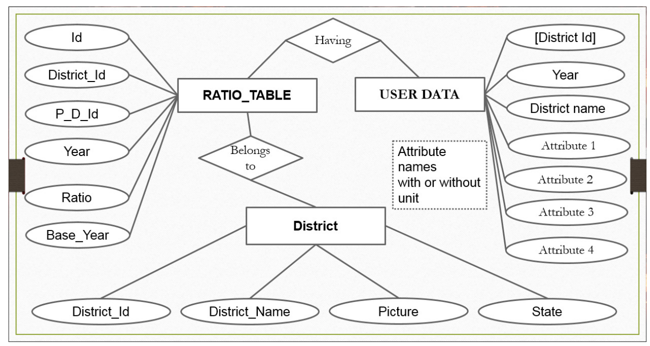


Fig. 1. ER-Diagram of the DAS application

Table 3. Modules in DAS

Module Name	Description
Registration	New user creation
Login	Login to registered users
Update Account	Change of password
Apportioning	Uploads the excel file,select the attributes, data apportioning
District Change	Provide information about district dynamism regarding changes in district boundary
District Code	Provide pre-defined district code for apportioning
Contact Us	Contact details
Help	Online help about software
Admin Home	Management of users
Data Management	Manage backend data request for apportioning temporal data of district and ratio of district boundaries stored by admin
Image	Administration regarding updating maps of the districts and other district details

2.2.2 Input Data Design

Data is required to be provided to the software using excel file. The columns of the excel file represent the attributes and the rows contain the dataset instances. Among the attributes District Id and year are mandatory. Remaining attributes are of either absolute or density category. Input data format is shown in the Table 4. The variables from 1 to n are for illustration purpose. Depending on user requirement of apportioning data, the variable can be defined.

Table 4. Input data for apportioning

Attribute	Description
District Id	Standard District code provided for each district from DAS Application
District Name	Respective district name of particular District id
Year	User can apportion the data for 3 decades from 1980-2013
Variable 1	Any absolute or density variable
Variable 2	Any absolute or density variable
Variable n	Any absolute or density variable

2.2.3 Output Design

Various outputs that user expects from a data apportioning software are

1. Report on selected district dynamics related to its area

Sometimes a user will be interested in knowing the trends of changes in area of a selected district. To meet this requirement, available static map of the district along with the list of changes in the history of the selected district should be made available (Table 5). The Fig. 2. shows the boundary of Madurai district.

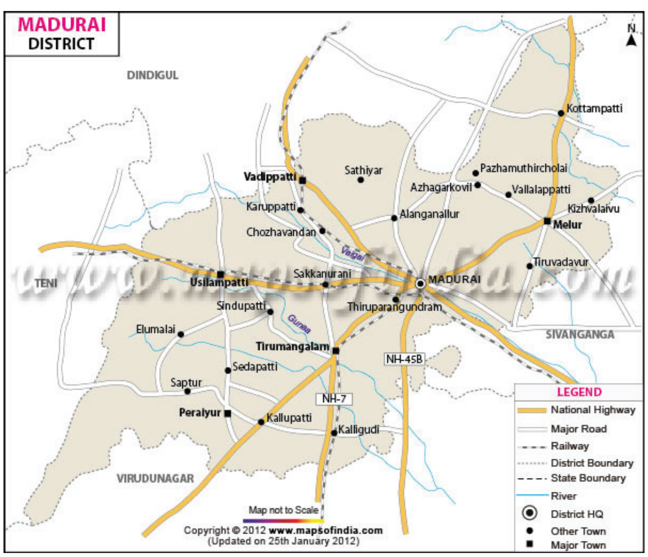


Fig. 2. Sample map to showing boundary of Madurai district

Table 5. Area dynamics of Madurai districts

ID	Parent-District	District name	Start-from	End-year	Area
8009	Madurai	Madurai	1980	1985	12623
8009	Madurai	Madurai	1986	1995	6565
8009	Madurai	Madurai	1996	2014	3696
8019	Madurai	Dindigul	1986	2014	6058
8028	Madurai	Theni	1996	2014	2869

2. List of Codes of the districts

District code is necessary for apportioning process. So standard district ids are needed to be appended in data file. To fufill this requirement, the district codes are provided through DAS. Table 6 shows the sample format of district codefor the state of Tamilnadu.

Table 6. Sample format of District Code

District_Id	District_Name
8001	The Nilgiris
8002	Kanyakumari
8003	Chennai
8004	Tirunelveli
8005	Vellore
8006	Cuddalore
8007	Ramanathapuram
8008	Coimbatore

3. Apportioned Data in Excel format

Apportioned data of selected attributes are presented in tabular format with the scrolling facility. The resulting apportioned data can be exported in excel format as shown in Table 7. Users can view and download the apportioned data for a desired base year. Results format are presented in Table 7.

Table 7. Output design Format

Attribute	Description
Id	Standard District Id of districts
District	Respective district name of particular District id
Year	Temporal data contain year from user data
Var1 to n	Apportioned absolute or density variable

3. FUNCTIONALITY OF DAS

Homepage of the DAS web based software is depicted in Fig. 3. DAS has been designed for two different user category (administrator and general user) with different access rights. After authentication general user gets access to apportioning, district dynamism menu item and other features of general use. Admin user gets access rights to management of user information, district details and data management of district boundaries changes. User has to be a registered member of the system.

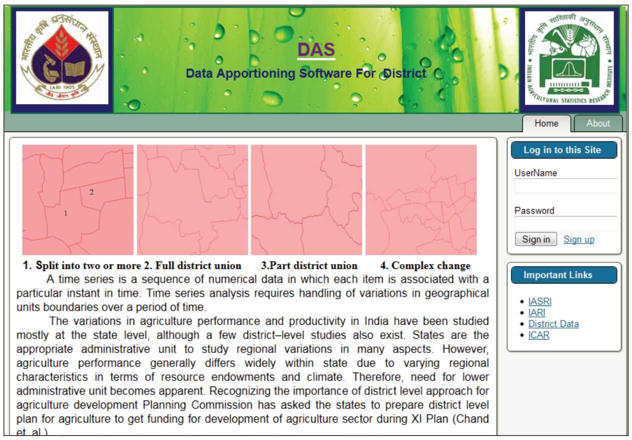


Fig. 3. Homepageof DAS

3.1 Selection of Attributes

Sometimes input data file may have some irrelevant attributes which are not to be used and requires manual filtering. The software helps in this process after uploading excel file data.

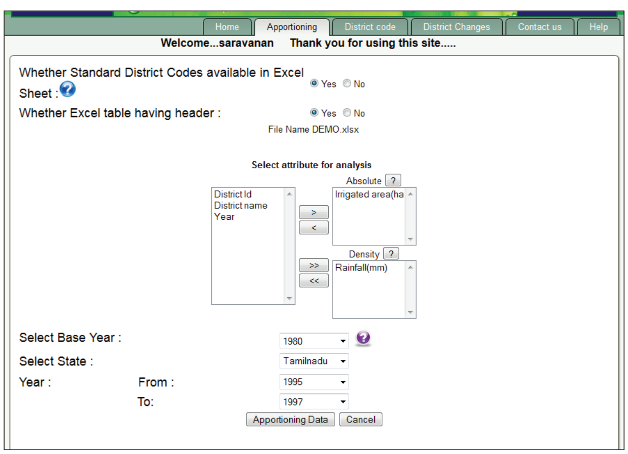
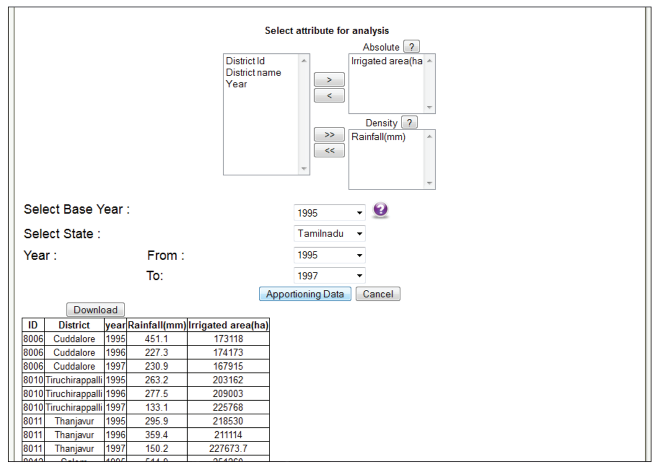


Fig. 4. Attributes selection in DAS

DAS application can work with both types of attributes for apportioning namely absolute attributes and density attributes. User can select the required attributes from the whole list of attributes available (Fig. 4). Two list boxes (Absolute and Density) are available where the user can transfer the desired attributes given in the left list to the right list.

3.2 Apportioning

After attribute selection, the base year needs to be selected as per requirement. Fig. 5 shows the apportioning of data for the Tamil Nadu districts. Based on the base year (1995 in the Fig. 5) and the user data, apportioning is done on clicking the apportioning data button and the apportioned data is shown instantaneously. This process is very time consuming when done manually. Further, in manual computation reliability of results needs to be ensured by repeating the calculations.



ID	District	year	Rainfall(mm)	Irrigated area(ha)
8006	Cuddalore	1995	451.1	173118
8006	Cuddalore	1996	227.3	174173
8006	Cuddalore	1997	230.9	167915
8010	Tiruchirappalli	1995	263.2	203162
8010	Tiruchirappalli	1996	277.5	209003
8010	Tiruchirappalli	1997	133.1	225768
8011	Thanjavur	1995	295.9	218530
8011	Thanjavur	1996	359.4	211114
8011	Thanjavur	1997	150.2	227673.7

Fig. 5. Apportioned Data output page in DAS

3.3 Maintenance

The maintenance of any software is virtually never ending phase. Generally, the problems which are not observed during the development life cycle emerge during its practical use. DAS provided Administration interface to maintain various data used for analysis and also user registration data to be stored in database. DAS being a web based software is accessible to administrators as well as clients through internet. Administrator has the highest privilege among all users. Admin can do any modification in any part of the data captured in the database. Therefore for security they can only access the system after entering authentic username and password. When the administrator enters the authentic username and password in login form of DAS home (Fig. 6), then he could use user management module which is developed for monitoring access of different users. It is worth mentioning here that the software requires apportioning information for the districts in the database at the backend in the ratio table. Admin user can insert or update or delete the data stored in the database (Fig. 7).

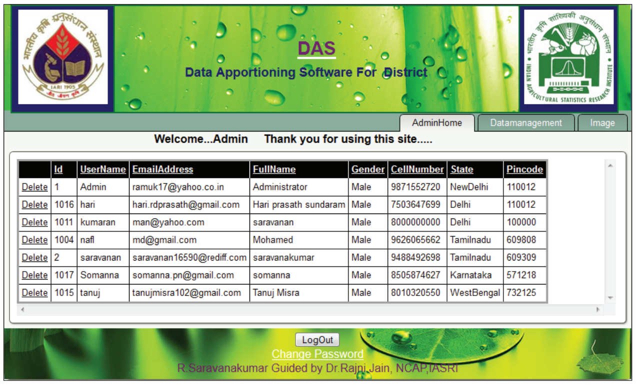


Fig. 6. User Management in DAS

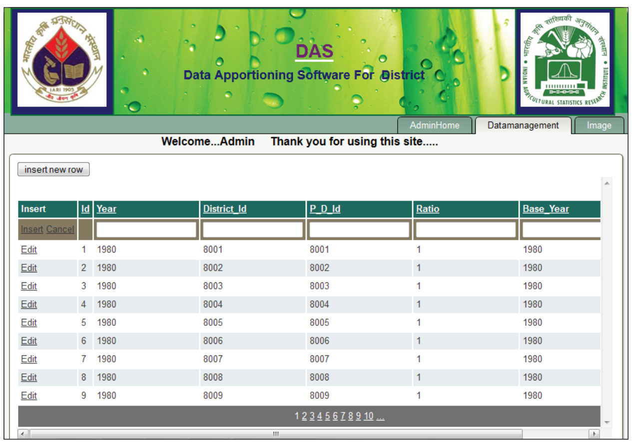


Fig. 7. Data Management in DAS

4. RESULT AND DISCUSSION

DAS meets the requirements that guided its design and development. However, DAS can be further enhanced by adding other methodologies of apportioning in future. Integrated software as well as unit level testing has been done for DAS using test dataset. Sample test dataset taken from [http://www.indiastat.com]. User can use more number of attributes to apportion at a same time. Test data and that result are shown in table 8, 9 and table 10.

Table 8. Sample test data

District name	District Id	Year	Irrigated area(ha)	Rainfall(mm)
Chengalpattu	8016	1995	331423	625.5
Tiruvallur	8025	1996	151519	346.6
Kanchipuram	8024	1996	184377	559.9
Tiruvallur	8025	1997	152822	322.2
Kanchipuram	8024	1997	187932	302.3

Table 9. Result Based on 1995 base year

ID	District	year	Rainfall (mm)	Irrigated area (ha)
8016	Chengalpattu	1995	625.5	331423
8016	Chengalpattu	1996	453.3	335896
8016	Chengalpattu	1997	312.3	340754

Table 10. Result Based on 1997 base year

ID	District	Year	Rainfall (mm)	Irrigated area (ha)
8024	Kanchipuram	1995	625.5	230829.5
8024	Kanchipuram	1996	559.9	184377
8024	Kanchipuram	1997	302.3	187932
8025	Tiruvallur	1995	625.5	100593.5
8025	Tiruvallur	1996	346.6	151519
8025	Tiruvallur	1997	322.2	152822

5. CONCLUSION

The paper presents knowledge engineering for apportioning of district level data with the help of mathematical notations. In general there are three kinds of changes (i) one district splits into two or more districts (ii) two districts combine to make a new district (iii) two or more adjoining old districts provide two or more new districts with changed boundaries. Online software DAS has been developed to replace manual or spreadsheet computation method for apportioning district level data. DAS is online and is compatible with any operating system. It allows the user to apportion both absolute and density type variables. It saves lot of time and make it comfortable to work with district level time series data. However, data which requires apportioning needs to be validated for this software by the administrator of the software. Online help facility is also available. Besides it provides information on district boundary dynamics along with its map. Data is validated by the system automatically. Output is available online with export option of apportioned output into Excel format for further analysis.

DAS is easy to use and is useful in saving time for apportioning district data. DAS is user friendly and no need of any expertise or computer programming to use it. Administrator interface of the software helps in development and maintenance of user database.More Advanced Techniques like Machine Learning along with Geo-referencing may be used in future as a more widespread coverage of the similar problems.

REFERENCES

Bhalla, G.S. and Alagh, Y.K. 1979. Performance of Indian Agriculture: A District Wise Study. New Delhi: Sterling, 2-8.

Bhalla, G.S. and Tyagi, D.S. 1989. Pattern of Indian Agricultural Development: A District Level Study, Institute for Studies in Industrial Development, New Delhi.

Bhalla, G.S. and Singh, G. 2001. Indian Agriculture: Four Decades of Development, New Delhi: Sage, 1-8 & 216-223.

Bhalla, G.S. and Singh, G. 2010. Growth of Indian Agriculture: A District level Study, Planning Commission of India Project Report, Jawaharlal Nehru University, New Delhi.

Chand, R., Garg, S. and Pandey, L. 2009.Regional Variations in Agricultural Productivity: A District Level Study. National Centre for Agricultural Economics and Policy Research, New Delhi.

Bos, B. and Lie, H.W. 1999. Cascading style sheets: Designing for the Web (2nd ed.). Harlow, England: Addison-Wesley.

http://www.indiastat.com

http://www.icrisat.org/vdsa/Include/Document/all-apportioned-web-document.pdf

http://www.mapofindia.com

http://vdsa.icrisat.ac.in/

Griffiths, I., Flanders, J. and Sells, C. (2003). Mastering Visual Studio .NET. Beijing: O'Reilly.

Grove, R.F. 2010. Web-Based Application Development. Sudbury MA: Jones and Bartlett Publishers.

Jain, R., Alam, A.K.M.S. and Arora, A. (2013). WBSTFP: Software for TFP Computation in Agriculture. *J. Ind. Soc. Agril. Statist.*, New Delhi.

Kumar, A. and Jain, R. 2013. Growth and Instability in Agricultural Productivity: A District Level Analysis, *Agrict. Economics Res. Review*, 26 (conf.), 31-42.

Kumar, H. and Rohini, S. 2009. Mapping Indian Districts across Census Years 1971-2001. *J. Development Economics*, Centre for Development Economics, New Delhi.

Munassar, N.M.A. and Govardhan, A. 2010. A Comparison between Five Models of Software Engineering. *Int. J. Comp. Sci. Inf. Tech.*, New Delhi, 94-101.

Macdonald, M. 2002. ASP .NET: The Complete Reference. Berkeley CA: Tata McGraw Hill.

Powell, T.A. 2003. HTML and XHTML: The Complete Reference. Berkeley CA: Tata McGraw Hill.

Sommerville, I. 2011. Software Engineering: A Practitioner's Approach. Boston: Pearson.