# DATA STRUCTURE MAPPING OF BLACK SOYBEAN MULTI-TEMPORAL AND MULTILATERAL OBSERVATION WITH GRAPHICAL MODELING

H. Umakawa [a, *], K. Sudo [b], K.K. Mishra [a], K. Tachibana [a], H. Shimamura [a]

[a] R&D Center, PASCO Corporation, Meguro, 2-8-10 Higashiyama, Meguro-ku, Tokyo 153-0043, Japan
hiroko_umakawa@pasco.co.jp
[b] Hyogo Prefectural Technology, Center for Agriculture, Forestry and Fisheries, Kasai, Hyogo 679-0198, Japan

**Commission VIII**

**KEY WORDS:** Black soybean, Multi-temporal, Agriculture, Water stress, Abstraction, Analytical, Visualization

**ABSTRACT:**

A semi-automatic visualization methodology was developed in this study for the structures of several relationships among many numerical datasets acquired for multi-temporal, multilateral and large-scale observations. The assumptions of this method are followings: (i) all inputs are evaluated by appropriate method for each dataset, and (ii) all input data is indicative to a state of a target object during the observation. The developed methodology in the current study can abstract relative relationships of dataset pairs as a data structure map data utilizing graphical approach. This leads to the advantage that any manual editing is admissible for modification of local model among datasets for improvement of graphical appearance, since the determined coefficient is adapted as a factor to dictate the relationship strength. Additionally this adoption makes processes simpler to fuse or deconstruct the nodes. A concept "Layer" makes structure rationalization of a map easier since it utilizes categorization according to data acquisition and observation techniques. Furthermore, load reducing of calculation resulted real-time map editing. We applied the data mapping method for three years of observations for the project of the most popular black soybeans "TANBA NO KUROMAME" in Hyogo prefecture and to gain an overview of the acquired database. The project obtained three years of yield data and temporal firsthand observation data, temporal spectral data which were acquired by five airborne hyperspectral, multi-spectral and broad-band sensors in nine times over three crop growth periods. Every map of several yield data showed a similar structure and we found that the most effective wavelength ranges for investigation items of the yield difference for individual growth stage through a composite map of yield data and aerial spectral reflectance data.

## 1. INTRODUCTION

### 1.1 Background

In the recent years in Japan, many farmers have high expectations for the development of agricultural management assistance systems utilizing remote sensing technology. Furthermore, agronomists are also concerned about adopting remote sensing technology for rapid acquisition of advanced information about crops and farm state over large farmland areas. In the future, we can anticipate the popularity of multi-temporal observations and organizational initiatives for various farming assistances

For a smooth implementation of any large observation project, we must enhance two necessary functionalities: (i) operational assistance to handle its huge database which is collected with multiple technology resources, and (ii) easy sharing of each interpretations and products among project members from various backgrounds.

### 1.2 Group and diagram technique

For improving such sophisticated technical functionalities, we can use some tools to assist communication and expression for better understanding of the data. Thinking is also a technique that used to be considered in the middle of the nineties and it has more algorithmically facilitated the process for problem solutions or a method of ideas through progressive era of Artificial Intelligence Technology (Buchanan, 2005) and recently it has become one of the most common tools. Brainstorming is another most popular group creativity technique to trigger an idea and chain reaction on an issue by bouncing ideas on each other in a small group. This method commonly uses organization and representation tool (such as, KJ method, concept mapping, mind mapping etc.) on its evaluation step because it usually produces barrage of information. Most of representation tools give us structure diagrams based on links among our ideas to order our thoughts better.

### 1.3 The hardship of biological modelling

In Japan, different kinds of crops are grown in a farm area. It is conceived that organized procedures of multi-period and multi-crop observations can reduce the survey cost. And it can also fulfil another demand to increase the crop varieties and number of samples. However the project will get a consideration so that participating members have to increase their workload in parallel for the topics, especially for the surveying. On the other hand, modelling ideas are produced through so many try-and-select process, and it takes a lot of efforts, costs and high expenses. Especially we need to have multi-sided approach to analyze relationships among various datasets. This study attempted to support such working processes by intuitive

---

* Corresponding author.

representation of relationships and automation of arithmetic efforts of modelling.

## 1.4 Black Soybeans

Black soybean is one of the most popular viands in Japan. A special Japanese black soybean, known as "TANBA NO KUROMAME", has been cultivated particularly in the western parts of Japan. The "TANBA", historical name of the region, stretches over Kyoto and Hyogo prefectures. This black soybean has been historically most famous brand among other Japanese black soybeans due to beauty of its black color, largeness, circular shape and good quality. Actually it is one of the most executive traditional cash crops. This name has been kept best by self discipline of producer's associations in the area. However, production control by the discipline increases the impact of yield reduction and quality deterioration of the seeds.

The control of soil moisture content considered as one of the most important field works in black soybean cultivation. In general, soybean yield and its seed quality are influenced by soil moisture during its whole growing process. Water stress after the onset of summer season is a main factor of bad podding, delay of harvest and green stem syndrome. The water stress affects the stable production (e.g., yield reduction and seed quality loss). Especially, black soybean requires a trained eye and good experience to judge whether the stock has been dried. Furthermore, soybeans fields are commonly rotated taking three or four calendar years due to easy harm by continuous cropping. The farmers cultivating black soybean need advanced skill to manage soil water content all over their fields depending on the conditions of particular stock, soil, weather and climate. Figure 1 shows black soybean plot.



Figure 1. Black soybean and its farm

## 1.5 Implementation of the development project to provide effective information for black soybean cultivation

The Ministry of Agriculture, Forestry and Fisheries of Japan is conducting a competitive research funding system, "Research and development projects for application in promoting new policy of Agriculture Forestry and Fisheries", for the industry-academia-government collaborative research groups. Under the funding scheme, a three-year project entitled, "Development of stable productive technologies of black soybean for field rotation in the KINKI region", had been conducted since 2007 for three consecutive years. Participants of the project achieved their own results for black soybean cultivation, namely: (i) it was investigated that water stress caused by soil desiccation deeply involved in pod setting, delayed harvesting and straight head, and (ii) a diagnostic procedure of spotting water supply point based on personal soil moisture meter was developed, (iii) a water balance model for black soybean field was developed based on solar radiation, atmospheric temperature and water supplying information. This project aimed to develop the procedure to judge and pin point the water stress area and best

timing for irrigation supply. Based on the above mentioned findings and airborne spectral detection technique, the project developed a system to diagnose water supply time locally by the farmers themselves. The system needs an estimation of local water balance with a personal soil moisture meter records, water supplying history and remotely sensed surface temperature of the cropped area. Furthermore the project is providing the management system for local farming assistance projects. The project members have already done promotions for the local farmers and agencies and introduced the system for the black soybean cultivation.

One of the interesting features of the project was to conduct airborne spectral observations and field surveys about soil moisture and growth, foliage temperature etc., in parallel at several crop stages during an enormously influential period "flowering – pod setting- seed filling" in each of the three years. This study applied a data structure abstraction methodology to a huge database of the project to gain an overview of data structure, and then we performed the validity check of the output results.

## 2. MATERIALS

The field surveys and mutli-temporal airborne spectral observations for three growth periods were carried out from 2007 to 2009.
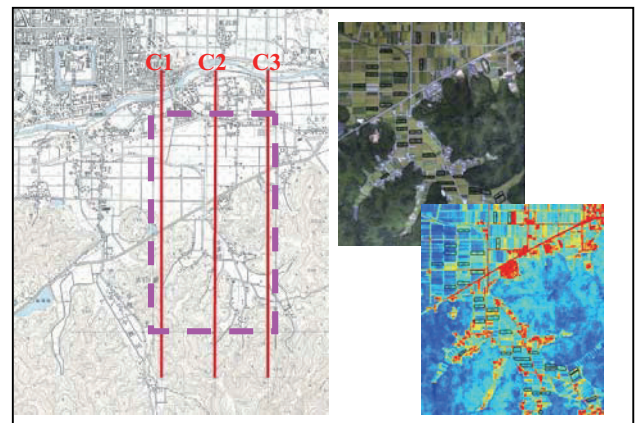


Figure 2. Test site (VIS and TIR ortho images)

Airborne campaigns with five sensors were conducted nine times for the observations of: (i) hyperspectral sensor AISA, (ii) multi-spectral sensors AZM and ADS40, (iii) broad-band thermal sensors TABI-320 and TS7302. The wavelength range of the data is ranged from visible to near- and mid-infrared, thermal infrared wavelengths. Figure 2 depicts the test site and airborne visible and thermal ortho images.

Field verification works in approximately 30 farms during the aerial observational days were also carried out. Three measurements points were set into each farm to check growth and soil moisture content. Main stages of the field surveys are as below: (i) flower bud differentiation – flowering stage: crop growth (main stem length, nodes, LAI, etc.), (ii) flowering - pod formation stage: soil moisture content (PF meter), leaf temperature, aperture of stomata, etc.). At end of the season, yields were observed including, number of pod, number of seed, seed size, seed weight, etc. The study verified the linkages between yield data versus spectral data.

### 3. DATA ABSTRACTION METHODOLOGY

#### 3.1 Overview

The methodology abstracts relative relationships of different pairs of datasets in a database and represents the utilization of Graph theory as a data structure map. The assumptions of this method are following: (i) all inputs are evaluated by appropriate method for each dataset, and (ii) all input data is reflective to a state of a target object when the observation is conducted.

#### 3.2 Configuration of components

**3.2.1 Node and Edge**: Every dataset is expressed as point (node) and a relationship among 2 datasets and is depicted as a line (edge) between the nodes. Length of edge is proportional to strength of the relationship between nodes that creates the edge. In particular, it shows how best the datasets can fit into a given model. A linear model with two variables is given for the relationship as an initial model. Initial edge length is derived from magnitude of a determination coefficient of the initial model (Figure 3).
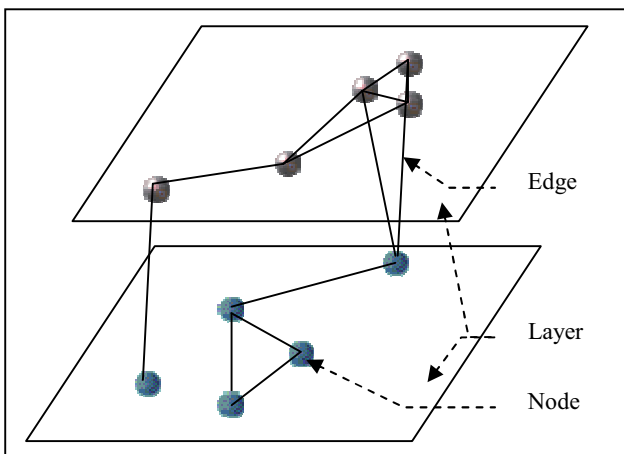


Figure 3. Conceptual diagram of different components

**3.2.2 Layer**: Added to the basic structure, we use a higher-level concept "Layer" to gather the result of every pro-temporal investigation. Every node on a layer are linked and adjusted with the heterogeneous nodes on another layer using same method as drawing inter-layer edges. The structure of map's graph is simplified after consolidation of similar links in several layers. Realistically multi-temporal field survey and hyperspectral observation provide enormous quantity of datasets into the database. For handling huge database, layer information is a helpful for tying graphical components on the data structure map. There is a good prospect to utilize the categorization by data characters (for example, data acquisition time, collection method, production technique etc.).

#### 3.3 Functions for relative editing

**3.3.1 Edge relaxation based on spring model**: By adjusting the position of the nodes and lengths of the edges which harmonize entire dataset, a complicated graph with many datasets is automatically processed, not to overlap each node and not to put figures too close. For this adjustment, a new function is added to the models of edge length and node positions, using spring model which is proportional to each determination coefficient. In particular, equilibrium of force over the wire frame is harmonized by the adjustment of spring length based on each determination coefficient.

**3.3.2 Multivariate relationship model**: Furthermore, to cope with multivariable model given, this method has only two calculations (center point position and determination coefficient of the model). The selected parts of graph which forms polygon shape are re-evaluated with new model and replaced its form to a kernel point.

**3.3.3 Advantages of the methodology**: Any manual editing is admissible for changing local model among nodes and for improving the appearance of the graph under this methodology. The method employs coefficient of determination to draw the graph, and user is able to freely change each local model according to their interpretations. Furthermore, this semi-automatic abstraction methodology has increased the efficiency of the overall proceedings and reduced the complex issues.

### 4. DATA STRUCTURE MAPPING OF DATABASE FOR BLACK SOYBEAN OBSERVATIONS

#### 4.1 Yield datasets

First of all, we applied the mapping method for several yield data. Some changes of the yields were noticed due to weather conditions during the three years of study period. However, by our graph (Figure 4), which was created with automated processing by initially adopting the linear model, we reviewed well-known relationships between several dataset of each year which were, for example, number of seeds – seed weight yield, number of pods – number of single seed pods, number of single seed pods - number of ripened pods etc. We reviewed well-known relationships between several dataset for each year which were for the number of seeds – seed weight yield, number of pods – number of single seed pods, number of single seed pods - number of ripened pods etc. Figure 5 shows segmentation of yield survey items.
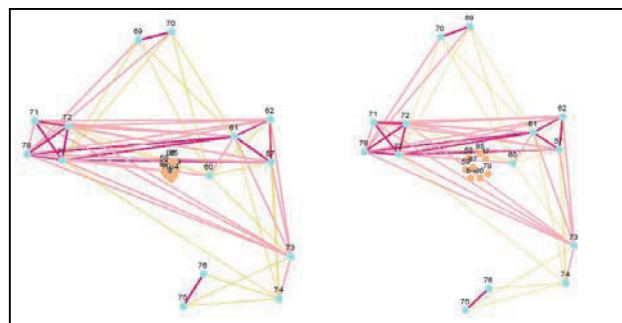


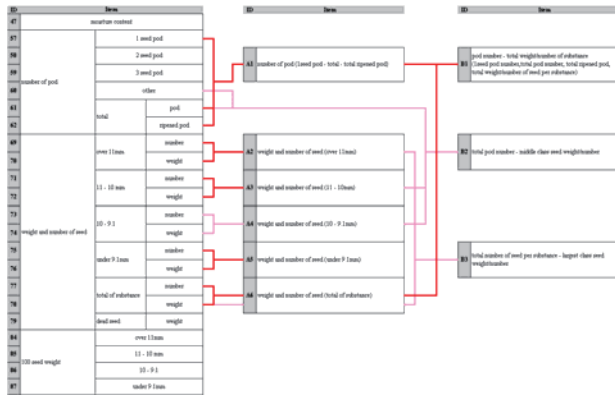Figure 4. Structure map of yield data (2008, 2007)

Figure 5. Segmentation of yield survey items

## 4.2 Spectral datasets

We conducted similar processing of the data captured by airborne spectral sensors and finally connected all layers onto the yield layer.

## 4.3 Composite mapping

Comparisons were conducted for the most effective wavelength range to each yield item at several growth stages. According to our created map, the most effective wavelength ranges for investigation items of the yield were noticed from changing throughout several growth stages (Figure 6, 7 and 8).
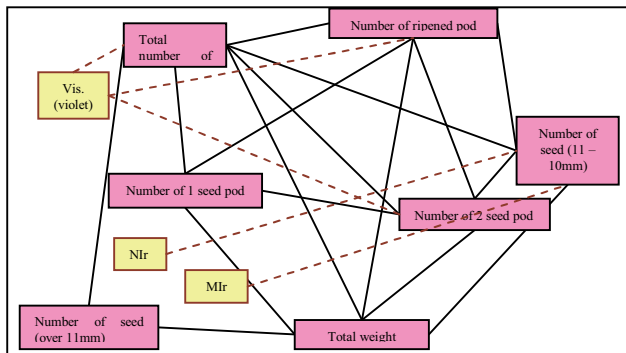


Figure 6. Effective wavelength regions (flower bud differentiation stage)
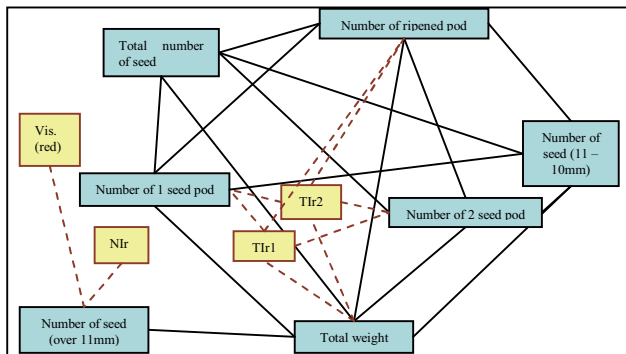


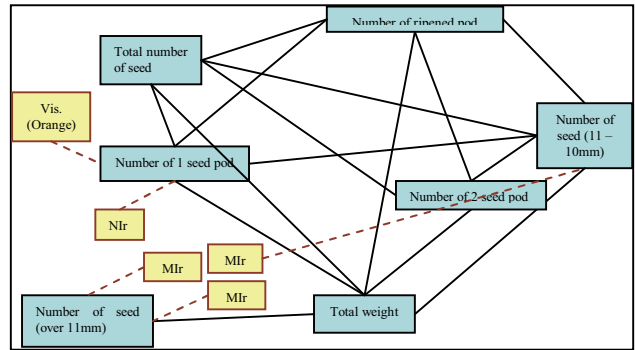Figure 7. Effective wavelength regions (pod filling stage1)



Figure 8. Effective wavelength regions (pod filling stage2)

## 5. CONCLUSIONS

This study presented a novel methodology to abstract multilateral structure of several relationships among various kinds of voluminous multi-temporal datasets, and confirmed the results by comparisons with the results and well-known relationships of the acquired observation data. This methodology reduces any kind of raised issues or confirmation for each dataset and its relationship. Furthermore, the created map indicated that the most effective wavelength ranges for investigation item of the yield differed in individual crop growth stages. Every map of several yield data showed a similar structure and we found that the most effective wavelength ranges for investigation items of the yield difference for individual growth stage through a composite map of yield data and aerial spectral reflectance data

**References:**

Bishop, C.M., 2006, Chapter 8. Graphical Models. Pattern Recognition and Machine Learning. Springer. pp. 359–422

Buchanan, B.G.., 2005, A (Very) Brief History of Artificial Intelligence, AI Magazine. pp. 53−60