

AUTOMATIC PARAMETER ESTIMATION IN MRF BASED SUPER RESOLUTION MAPPING

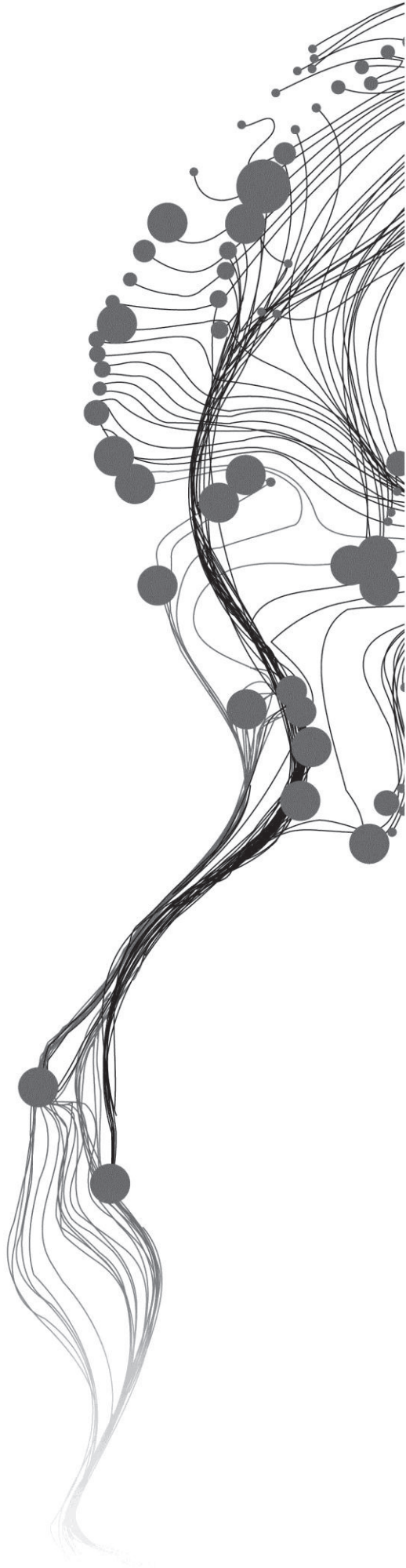
ANTENEH LEMMI ESHETE

March, 2011

SUPERVISORS:

Dr. V. A . Tolpekin

Dr . N. A. S. Hamm



AUTOMATIC PARAMETER ESTIMATION IN MRF BASED SUPER RESOLUTION MAPPING

ANTENEH LEMMI ESHETE

Enschede, The Netherlands, March, 2011

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.

Specialization: Geo-informatics

SUPERVISORS:

Dr. V. A. Tolpekin

Dr. N. A. S. Hamm

THESIS ASSESSMENT BOARD:

Prof. Dr .lr. A. stein (Chair)

Dr. Ir .B .G .H .Gorte (External Examiner)

DISCLAIMER

This document describes work undertaken as part of a programme of study at the Faculty of Geo-Information Science and Earth Observation of the University of Twente. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the Faculty.

ABSTRACT

In Markov Random Field (MRF) based super resolution mapping (SRM) the accuracy of classification is depend on the optimal parameters. The smoothness parameter λ balances the contribution of the prior and likelihood energy terms. Whereas, λ_p parameter balances the contribution of likelihood energy from the panchromatic and multispectral band. By proper setting of these parameters good classification accuracy can be obtained. However, poor parameter setting produces unsatisfactory results. Trial and error estimation of the parameters is time consuming. Therefore, this study concentrate on developing new models to estimate the optimal smoothness parameter λ and λ_p parameters based on local energy balance analysis. The study shows how the optimal values of the parameters depend on the scale factor and class separability information.

The data sets used during this study were synthetic images, generated systematically with various class-separability values. This enables to evaluate the models at different class separability and scale factor information. The contextual and spectral information were modelled with prior and the likelihood energy functions respectively. The global energy was constructed and different combination of λ and λ_p parameters were tried. To find the minimum of the total energy in map estimate simulated annealing algorithm was used. An optimal λ and λ_p values were identified based on kappa value and in order to test the predicted λ and λ_p values a range for the optimal λ and λ_p value was specified.

An optimal value of λ and λ_p parameters exists for each combination of scale and class separability values in the panchromatic and multispectral image. The result obtained from the developed model for the optimal λ and λ_p parameters agree with the empirical data. The study shows that the incorporation of information from the panchromatic image increases the classification accuracy at the lower scale factors, if it is properly estimated.

Key words

Class separability, Markov random field (MRF), Super resolution mapping (SRM).

ACKNOWLEDGEMENTS

I express my special gratitude to my supervisor Dr. V. A. Tolpekin for his advice, help and continuous encouragement have contributed significantly to the completion of this study. I want to extend my appreciation to my second supervisor: Dr. N. A. S. Hamm for his critical comments and advice that enhance this work.

A would like to thank Juan Pablo Ardila Lopez, a PhD student at earth Observation Department of ITC faculty, University of Twente for his advices and supports.

I am very grateful to my GFM 2010 colleagues for their friendship over the last 18 months study period. I am thankful to my Ethiopian colleagues for the company and moral support.

Enumerable thanks to my family members for their love and support during my studies in the Netherlands. This work would not have happened without their advices and supports.

Above all, I express my greatest thanks to Almighty God, who made me His creature and gave me His divine grace to successfully accomplish this study; all this would not have been possible without His will. Lord, Thank you

TABLE OF CONTENTS

1.	Introduction.....	1
1.1.	Background.....	1
1.2.	Problem statement	2
1.3.	Research objective.....	2
1.4.	Research questions	2
1.5.	Research setup	3
1.6.	Structure of the thesis	3
2.	literature review.....	5
2.1.	Previous works of MRF based super resolution mapping.....	5
2.2.	Parameter estimation techniques	6
3.	Methods.....	9
3.1.	Super resolution mapping(SRM)	9
3.2.	Synthetic data sets and class separability	10
3.3.	MRF	15
3.4.	Estimation of λ and λp parameter.....	17
3.5.	Simulated Annealing	20
3.6.	Accuracy assessment and performance analysis.	21
4.	Results.....	23
4.1.	Experimental results from synthetic datasets	23
4.2.	λ and λp estimation results from synthetic image	28
4.3.	parameter estimation result from the real image	32
4.4.	Summary of observation from the results.....	32
5.	discussions,conclusion and recommendation	33
5.1.	Discussions	33
5.2.	Conclusion	34
5.3.	Recommendations	34
	List of references	36
	Appendix.....	37
	A. Summary of results for optimal λ and λ pan values.	37
	B. Summary of the result for optimal λ and λ pan estimation with average class separability.	41

LIST OF FIGURES

Figure 3.1: Construction of synthetic images. (a) Fine-resolution multispectral image x . (b) Reference image with three classes, green: shadow vegetation, white: grass, brown: tree. 10

Figure 3.2: Degraded synthetic multispectral (30x30 pixel) and panchromatic (60x60 pixel) images. 11

Figure 4.1: super resolution map result.(a) lowest kappa, (b) highest kappa, (c) reference image. 24

Figure 4.2: Optimal λ value at different scale and class separability in the multispectral and panchromatic image. Lines are added to facilitate the interpretation of the data. 24

Figure 4.3: optimal λ_{pan} value at different scale and class separability in the multispectral and panchromatic image. Lines are added to facilitate the interpretation of the data. 25

Figure 4.4: The effect of class separability on the optimal values of λ and λ_{pan} . (a) and (c) show the change of the optimal λ and λ_{pan} values respectively, when class separability in the multispectral image changes for a fixed JMZ value of 0.02. Here, (b) and (c) show the change of the optimal λ and λ_{pan} values respectively, when class separability in the panchromatic image changes for a fixed JMY value of 2.0. 26

Figure 4.5: The effect of scale factor and class separability in the classification accuracy kappa value. 27

Figure 4.6: Optimal λ values for varying parameters S and JMY. λ' : The experimentally determined optimal values. λ_{range} : The range of λ corresponding to $K \geq 0.85 k_{max}$. λ^* : The estimation from the model. 29

Figure 4.7: Optimal λ_{pan} values for varying parameters S and JMY. λ'_{pan} : The experimentally determined optimal values. $\lambda_{pan range}$: The range of λ_{pan} corresponding to $K \geq 0.85 K_{max}$. λ^*_{pan} : The estimation from the model. 30

Figure 4.8: Optimal values for varying parameters S and JMY. λ'_{pan} & λ' : The experimentally determined optimal values. λ^*_{pan} & λ^* : The estimation from the model. $\lambda_{pan range}$ & λ_{range} : The range of λ_{pan} and corresponding to $k \geq 0.85 k_{max}$. Here (a) and (b) for JMY value of 0.5 and (c) and (d) for JMY value of 1.0. (e) and (f) for JMY value of 1.9 for a fixed JMZ value of 0.02. 31

LIST OF TABLES

Table 3.1: Jeffries-Matusita distance between the classes for $JM^{(y)} = 0.5$	13
Table 3.2: Jeffries-Matusita distance between the classes for $JM^{(z)} = 0.02$	13
Table 3.3: Relation between minimal and average Jeffries-Matusita distance in image y and z	13

LIST OF VARIABLES

$\alpha\beta$ Spectral class α and β

μ Class mean vector

A^t Transpose of the matrix A

C_α^{-1} Covariance inverse of spectral class α

$\text{Tr}(A)$ denotes sum of the diagonal elements of the matrix A.

C_α Covariance matrix of the spectral class α

σ_α^2 Variance of the spectral class α

avg average

B Bhattacharyya distance

JM Jeffries-Matusita distance

x Fine resolution multispectral image

λ_p (λ_{pan}) λ Panchromatic

$B_{\alpha\beta}^{(y)}$ Bhattacharyya distance in the multispectral image

$B_{\alpha\beta}^{(z)}$ Bhattacharyya distance in the panchromatic image

$JM^{(y)}$ Jeffries-Matusita distance in the multispectral image

$JM^{(z)}$ Jeffries-Matusita distance in the panchromatic image

$JM_{avg}^{(y)}$ Average Jeffries-Matusita distance in the multispectral image

$JM_{avg}^{(z)}$ Average Jeffries-Matusita distance in the panchromatic image

1. INTRODUCTION

1.1. Background

Land cover mapping is a typical and important application of remote sensing data. Accurate land cover information is needed at a reasonable cost to many planning and monitoring programs. Different organization like geological survey and national mapping agencies are interested in extracting land cover information from remote sensing images. Different image processing and analysis techniques are used to extract information from the image. The traditional approach to land cover mapping is through hard classification from remotely sensed data in which each pixel is assigned to one land cover type. However this classification method is not appropriate when land cover information is required at sub-pixel level particularly in coarse spatial resolution image, where the presence of more than one type of land cover classes in a pixel, which is commonly referred to as mixed pixel. In sub pixel classification the pixel is resolved in to various class proportions to solve the problem of mixed pixels. However, it does not account the spatial distribution of class proportions within the pixel (Kasetkasem et al., 2005).

To overcome these problems super resolution mapping (SRM) techniques have been developed. SRM is a land cover classification technique that produces maps of a finer spatial resolution than that of an input image. It can be considered as a further step after sub-pixel classification in a sense that not only the fractions of classes within coarse resolution pixels are estimated, but also the spatial distribution of class proportions within and between pixels is considered (Tolpekin and Stein, 2009). The class proportion in coarse resolution pixels is computed in the soft classification step with the use of techniques such as linear spectral unmixing, neural network and fuzzy classification.

Accuracy of any land cover classification technique is influenced by spectral separability of classes (Swain and Davis, 1978), which is the measure of similarity between spectral signatures. For low class separability, classification leads to confusion between classes. In SRM accuracy of classification is influenced by scale factor and class separability but it is difficult to separate the influences of the two, because class separability depends on class spectral variation, which in turn, depends on scale. (Tolpekin and Stein, 2009)

For many applications the information obtained from the single image is incomplete, imprecise or inconsistent. Additional source may provide complementary information (Solberg et al., 1996). Fine spatial resolution images provide fewer spectral bands and larger within-class variation than the coarser resolution images whereas, the spatial resolution of the coarser resolution images are less than the fine resolution images. Due to these reasons, techniques have been developed that make an advantage to use both fine spatial resolution and spectrally rich coarse spatial resolution images, such as image fusion. (Solberg et al., 1996)

Markov random field theory provide a suitable and consistent way for modelling context dependent entities such as image pixels and correlated features (Li, 2001). In SRM spatial dependency between and within pixels can be enhanced by integrating spatial context. Spatial context is defined by the correlation between spatially adjacent pixels in spatially neighbouring pixels. The spatial context is very important for

the interpretation of a scene and it may be derived from spatial, spectral and temporal attributes. More information can be derived by considering the pixel in context with other measurement and the suitable use of context allows the elimination of possible ambiguities, the recovery of missing information, and the correction of errors (Tso and Mather, 2001).

1.2. Problem statement

Markov Random Field model plays an important role in image analysis because it integrates the contextual information associated with the image data in the analysis process, through the definition of suitable energy functions. However, an MRF model usually requires an estimation of one or more internal parameters before the application of the model. Particularly in the context of supervised classification, trial-and-error procedures are typically used to choose suitable values for at least some of the model parameter(Solberg et al., 1996). MRF parameter estimation is difficult, especially when the number of information sources increase.

$$U(c | y, z) = \lambda U(c) + (1 - \lambda)(\lambda_p U(z | c) + (1 - \lambda_p)U(y | c)) \quad (1.1)$$

The model (1.1) was introduced by Tolpekin, et al. (2010). λ_p is an internal parameter of the MRF based SRM model which balances the contributions of the conditional energy from the multispectral and panchromatic image. Whereas, the smoothness parameter λ balances the contribution of the prior energy and conditional energy to the global energy. To increase the classification accuracy these parameters must be properly estimated. Trial and error estimation of these parameters is time consuming and inefficient. There is no existing method that estimates the optimal λ_p parameter automatically. To obtain higher classification accuracy efficiently both the λ and λ_p parameter must be estimated optimally. Therefore, this study focuses on developing a model for the automatic estimation of λ and λ_p parameters.

1.3. Research objective

1.3.1. General objective

The general objective of this research project is to develop models for the automatic estimation of the optimal parameters λ and λ_p in MRF based super resolution mapping.

1.3.2. Specific objective

To achieve the general objective the following specific objectives have been identified

- a. To identify how the scale factor and class separability affects the optimal λ and λ_p parameter estimation.
- b. To develop models to be used for the optimal λ and λ_p parameter estimation.
- c. To test the models.

1.4. Research questions

- a. How does the scale factor influence the optimal λ and λ_p parameters?
- b. How does class separability affect the optimal λ and λ_p parameters?
- c. How to develop models to be used for the optimal λ and λ_p parameters estimation?
- d. How should the models be assessed?

1.5. Research setup

The adopted setup is carried out in six phases:

1.5.1. Synthetic image generation

Several synthetic images are generated by varying the class separability in the assumed high resolution multispectral image.

1.5.2. Image degradation

The multispectral and panchromatic images are degraded spatially and spectrally from the assumed high resolution multispectral image by a degradation model.

1.5.3. Modelling the prior and conditional energies

The prior information is modelled from the assumed fine resolution multispectral image using the Markov Random Field model. The conditional energy of the panchromatic and multispectral image is modelled from the two spatially and spectrally degraded images.

1.5.4. Parameter estimation

The models are developed for the automatic estimation of the optimal λ and λ_p parameter in MRF based super resolution mapping based on local energy balance analysis.

1.5.5. Simulated Annealing optimization

The total conditional energy and the prior energy are integrated with the MRF model under the Bayesian framework to obtain the maximum probability that is used to get optimal solution for the output image. The MAP estimate for the super resolution map is found by minimization of the total energy by using simulating annealing algorithm.

1.5.6. Accuracy assessment and performance analysis

The accuracy is measured in kappa coefficient and the performance of the models is assessed with the numerical optimal value obtained by the experiment.

1.6. Structure of the thesis

This thesis contains five chapters. Chapter 1 describes the background, the problem statement, the objectives, the research questions and the approach of the research. Chapter 2 presents a literature review on previous work on MRF based super resolution mapping and parameter estimation techniques. Chapter 3 describes about synthetic image generation and λ and λ_p parameter estimation techniques developed in this study. Chapter 4 presents the result of the research. Finally Chapter 5 presents discussion, conclusion and recommendation for further research.

2. LITERATURE REVIEW

This chapter explains previous works of MRF based super resolution mapping and reviews existing parameter estimation techniques

2.1. Previous works of MRF based super resolution mapping

Kasetkasem, et al.(2005) introduced MRF model based approach for the generation of super resolution land cover map from remote sensing data. The method works based on the assumptions that there is no mixed pixel in the fine spatial resolution images and the spectral values of classes in a fine spatial resolution image follow a multivariate normal distribution. It was applied into two phases: in the first phase initial SRM is generated from fraction images and in the second phase optimized SRM is produced by updating the pixels iteratively. Before implementing the second phase, it is important to determine neighbourhood window size that influences the labelling of the central pixel in the optimization process. Some of the limitation of the method was the weights given to the neighbouring pixels were estimated from the ground truth data which is not always easy to obtain. Moreover the neighbourhood size was fixed to second order for any scale factor value which limits the effectiveness of the method to work correctly at any scale factor.

Hailu (2006) assessed the suitability of MRF based SRM techniques for super resolution land cover mapping with synthetic images and remotely sensing data. First the spatial and the spectral information were modelled with the prior and likelihood energy function then the smoothness parameter λ was introduced to control the two energy function. SA was used to perform global energy minimization and the result of the MRF based SRM was evaluated by comparing to the fine resolution reference map.

Hailu identified several factors that can affect the accuracy of SRM like the smoothness parameter λ , neighborhood size, initial temperature, temperature updating, and class separability, object size and scale factor. Finally, Hailu (2006) found that MRF based SRM method produces a high quality SRM when the neighborhood size grows in relation to the scale factor and the optimal value of the smoothness parameter was affected with the type of scene and class separability. The study also shows that, it is possible to get a reasonable accuracy even for poorly separable classes by setting the smoothness parameter to the optimal value.

Tolpekin, et al (2010) extended the contextual MRF based SRM method developed earlier for multispectral image to include the panchromatic band for individual tree crown objects extraction purpose in urban area. Because of the limited spectral information offered by the sensors, It is difficult to discriminate tree crown objects from other land cover classes such as grass and shrubs by using spectral pixel-based classification techniques. However, the problem was solved in this method using contextual classification approach and the SRM technique was used to solve problems related to spatial resolution of the sensors.

2.2. Parameter estimation techniques

A number of parameter estimation techniques for MRF have been developed by many authors. Serpico & Moser (2006) employed a Ho-Kashyap optimization for determining the parameters. Whereas, Jia & Richards (2008) presented a method for determining the appropriate weighting of the spectral and spatial contributions in the MRF based approach of contextual classification. Tolpekin and Stein (2009) used local energy balance analysis as a means for estimating the smoothness parameter in MRF model.

The determination of the MRF model parameter that weight the energy functions is a difficult issue and it is known that the performance of the model is dependent both on its functional form and on the accuracy of the model parameters estimation. Serpico and Moser (2006) presented an automatic supervised procedure for the optimization of the weight parameters of the combinations of distinct energy contributions in MRF models with the Ho-Kashyap algorithm. The Ho-Kashyap Algorithm is used for the optimization of the weight parameter setting involved with MRF models for supervised image classification.

The method uses training data in order to select a set of parameters that maximizes the classification accuracy by developing the linear relation between the energy function and the parameter vector. The method does not estimate the true values of the parameter of the MRF model. Instead, the parameter values giving the highest classification accuracy are searched. The basic idea of the method is to use the training set in order to state a condition of correct classification of the training sample and find a parameter vector that fulfils the condition.

To solve the parameter estimation problem, first the energy functions must be expressed as a linear combination of distinct energy contributions then the parameter estimation problem is expressed as the solution of a system of linear equalities. In this method ICM is used for the energy minimization process and it is initialized with a given label vector generated by a non-contextual supervised classifier and iteratively modifies the class labels in order to decrease the energy function. The HK-based parameter setting methods is coupled with the ICM classification approach for automatic contextual supervised classification. The major steps of HK-ICM method are

1. Generate an initial non-contextual classification map by using supervised classification.
2. Compute the energy difference matrix according to the MRF model and to the label vector.
3. Compute an optimal parameter vector by running the HK procedure up to convergence.
4. Generate a contextual classification map by running ICM up to convergence.

The method has been presented in conjunction with the ICM algorithm for energy minimization and it can be used in conjunction with other energy minimization techniques such as simulated annealing. The method overcomes trial and error parameter optimization problems in MRF based supervised image classification.

Jia & Richards (2008) developed a method that determines the appropriate weighting of the spatial and spectral contribution in the Markov random field based approach of contextual classification. In this method, first the spatial and spectral components are normalized to be in the range (0, 1) then the appropriate value for the weighting coefficient can be determined. The spatial information is incorporated in to the classification by changing the discriminant function through the addition of a term that distinguishes spatial correlations.

Tolpekin and Stein (2009) developed MRF parameter estimation method based on local energy balance analysis. In this method the optimal smoothness parameter is estimated by considering class separability and scale factor information. When the label of the fine resolution pixel $a_{j|i}$ is changed from the true class value $c(a_{j|i}) = \alpha$, to another wrong class value $c(a_{j|i}) = \beta$, the resulting change of the local prior energy will be

$$\Delta U_{\alpha\beta}^p = q \left| \sum_{l \in N(a_{j|i})} \phi(a_l) [I(\beta, c(a_l)) - I(\alpha, c(a_l))] \right| = q\gamma \quad (2.1)$$

According to Tolpekin and Stein (2009), The value of γ depends on the neighbourhood system size, the configuration of class labels $c(a_l)$ in the neighbourhood $N(a_{j|i})$ and the choice of power-law index n .

The smoothness parameter λ is defined as

$$\lambda = \frac{q}{(1+q)} \quad (2.2)$$

Where $0 \leq q < \infty$ is used to control the overall magnitude of the weights.

The change of the fine resolution pixel label $c(a_{j|i})$ from α to β also causes the change of the composition of the coarse-resolution pixel b_i ; hence a change in the mean vector and the covariance matrix of that pixel value. For equal covariance matrices for the class α and β , The change in local likelihood energy is expressed as:

$$\Delta U_{\alpha\beta}^l = \frac{1}{2} \left(\frac{\mu_\beta - \mu_\alpha}{S^2} \right)^t \left(\frac{C_\alpha}{S^2} \right)^{-1} \left(\frac{\mu_\beta - \mu_\alpha}{S^2} \right) \quad (2.3)$$

For equal covariance matrices for the classes α and β the B distance is equal to

$$B_{\alpha\beta}^{(y)} = \frac{1}{8} (\mu_\beta - \mu_\alpha)^t C_\alpha^{-1} (\mu_\beta - \mu_\alpha) \quad (2.4)$$

From the above equations, the change in local likelihood energy is expressed with divergence and scale factor as:

$$\Delta U_{\alpha\beta}^l = \frac{4B_{\alpha\beta}^{(y)}}{S^2} \quad (2.5)$$

The local contribution to the posterior energy from the considered pixel is lower for $c(a_{ji}) = \alpha$ than for $c(a_{ji}) = \beta$ and the contribution of the likelihood energy should compensate the gain in the prior energy. Thus the λ value can be determined from the balance between the changes in the prior and the likelihood energy values.

$$\Delta U_{\alpha\beta}^p = \Delta U_{\alpha\beta}^l \tag{2.6}$$

Solving for λ

$$\lambda = \frac{1}{S^2 + \frac{4B_{\alpha\beta}^{(y)}\gamma}{1}} \tag{2.7}$$

If the smoothness parameter λ is greater than the optimal value, the model will lead to over smoothing on the contrary λ value that is too small does not exploit the prior information in the model.

Having reviewed the well known parameter estimation techniques, the focus of this research is to develop models that estimate the smoothness parameter λ and λ_p in MRF based super resolution mapping.

The models developed in this research is similar to the method of (Tolpekin and Stein, 2009) and in both cases parameter estimation is done by considering class separability information between spectral classes. The detail discussion about the method will be given in chapter 3.

3. METHODS

The powerful property of the MRF models is that the prior information and the observed data from different sources can be easily integrated through the use of suitable energy functions. However, an MRF model usually requires an estimation of one or more internal parameters before the application of the model (Serpico and Moser, 2006). An appropriate choice of parameters can give successful result. Conversely, improper selection of parameter values will produce unsatisfactory result (Li, 2001). MRF Parameter estimation is difficult, especially when the number of information sources increases that make the number of parameter to be estimated increases.

The main focus of this chapter is to describe the MRF parameter estimation technique which is employed in this research. Section 3.1 describes about super resolution mapping. Section 3.2 explains how the synthetic images are constructed. Section 3.3 illustrates how the prior and the conditional energies are modelled. Section 3.4 briefly describes about the proposed method. Section 3.5 explains about simulated annealing optimization algorithm. Section 3.6 explains how the model is assessed.

3.1. Super resolution mapping(SRM)

SRM theory including its formulation described below is adapted from (Tolpekin et al., 2010) with some minor changes.

Consider the classification of a coarse resolution multispectral remote sensing image y that consists of K spectral bands with spatial resolution R and pixel locations $d_i \in D$, where D is $M_1 \times M_2$ pixel matrix and a panchromatic image z with finer spatial resolution $r < R$. The resulting super resolution map (SR map) c is a classified map which is defined on the set of pixel locations A and covers the same extent on the ground as y and z with spatial resolution r . The scale factor is denoted as $S = R/r$, which is the ratio between the coarse and fine resolution pixel sizes. Hence each pixel d_i will contain S^2 fine resolution pixels of a_{jli} .

Further assume a multispectral image x having the same spectral bands as y as well as the spatial resolution of r is defined on the set of pixels A . Image x is not observed directly while image y and z are obtained by spatially and spectrally degrading the image x . Furthermore it is assumed that every pixel in image x can be assigned to a unique class: $c(a_{jli}) = \alpha$, where $\alpha \in 1, 2, \dots, L$. The relationship between y and x , and z and x are established by a degradation model as:

$$y_k(d_i) = \frac{1}{S^2} \sum_{j=1}^{S^2} x_k(a_{jli}), k = 1, \dots, K \quad (3.1)$$

$$z(a_{jli}) = \frac{1}{K} \sum_k x_k(a_{jli}) \quad (3.2)$$

For $d_i \in D$ and $a_{jli} \in A$

3.2. Synthetic data sets and class separability

Synthetic image provides a useful source of data for improving our understanding of information extraction from remotely sensed data (Tatem et al., 2002). Tolpekin and Stein (2009) showed that the optimal smoothness parameter λ is dependent on the scale factor and class separability. Based on their findings, in this study it is assumed that the optimal λ and λ_p parameters depends on scale factor and class separability. The main advantage of using a synthetic image in this research is to explore the effects of scale factor and class separability on the optimal λ and λ_p parameters. The synthetic image allows to concentrate on specific element of the problem by ignoring the complexity of real images and used to perform systematic controlled parameter variation. (Tolpekin and Stein, 2009)

The synthetic image generation starts with the reference map (60×60) with three classes as shown in Figure 3.1b. Pixel values are generated through multivariate pseudo random number generator using class parameter mean and covariance. During image generation the class separability is controlled by fixing class mean and covariance matrices.

The fine-resolution multispectral image x Figure 3.1(a) is produced by sampling from multivariate normal distribution using mean and covariance obtained from the real image. This image is consequently degraded to coarse resolution multispectral and fine resolution panchromatic image with different scale factor values $S=2, 4, 5, 6, 8, 10$. (See equation 3.1 and 3.2). An example of degraded synthetic image with $S=2$ is shown in figure 3.2. Several fine resolution multispectral images x are constructed from the reference image with various class-separability values.

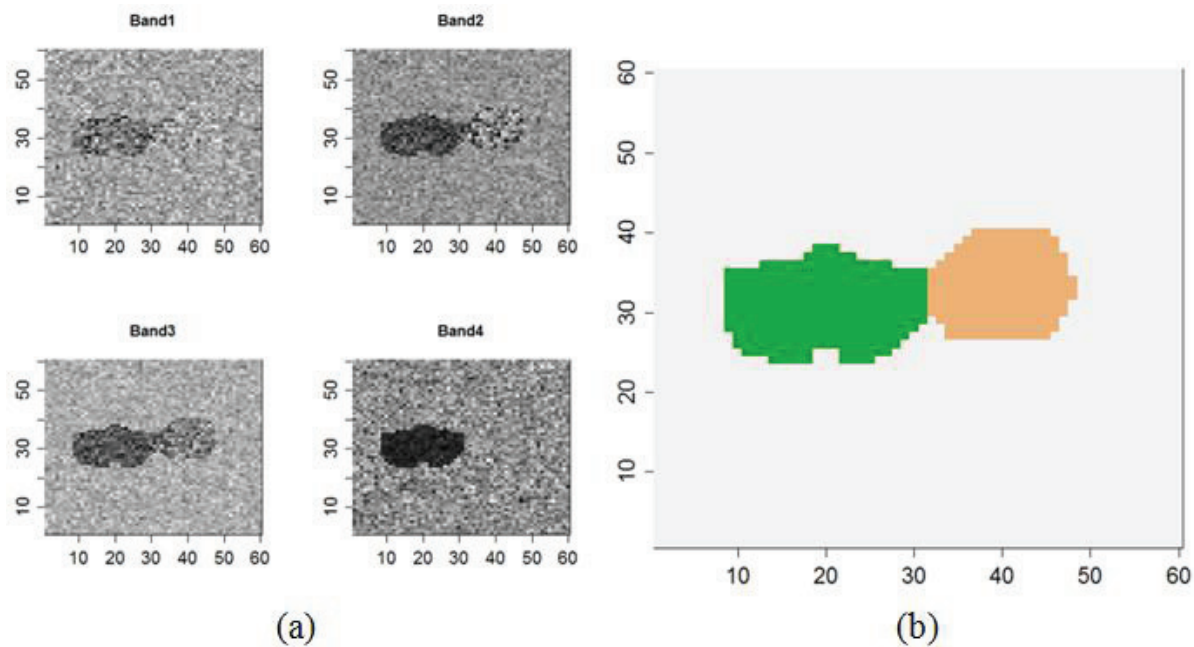


Figure 3.1: Construction of synthetic images. (a) Fine-resolution multispectral image x . (b) Reference image with three classes, green: shadow vegetation, white: grass, brown: tree.

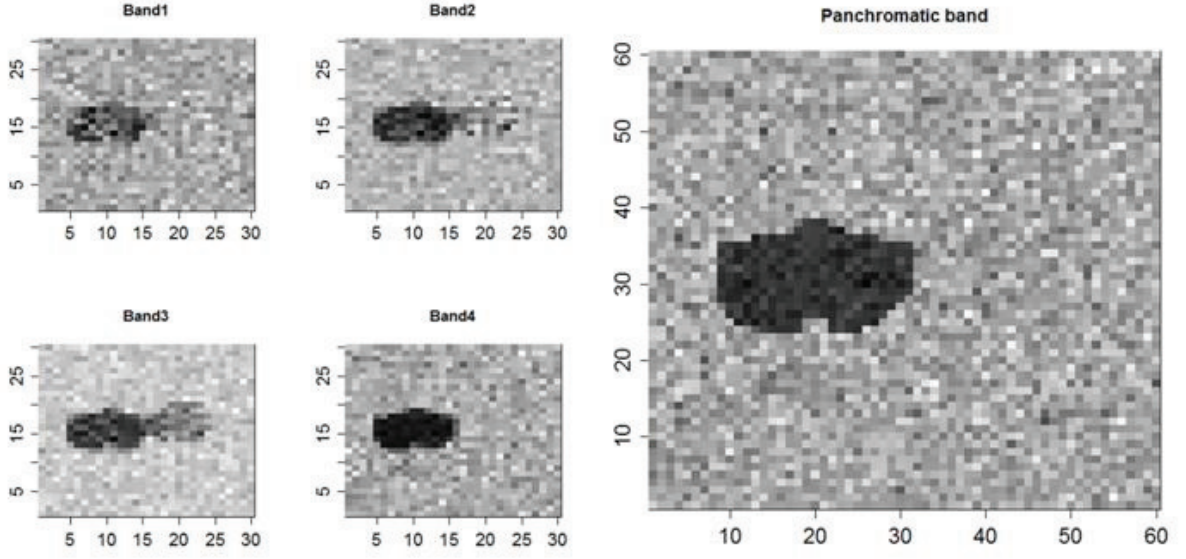


Figure 3.2: Degraded synthetic multispectral (30x30 pixel) and panchromatic (60x60 pixel) images.

Class separability is a statistical measure that shows how well the user defined classes can be separated by classifier. The simplest class separability measure is the Euclidean distance evaluation where the spectral distance between the mean vectors of each pair of class signature is computed, and if this distance is not significant for any pair of bands available they may not be distinct enough to produce successful classification. The basic principle is that pixel values within a given land cover type should be close together in the measurement space; whereas pixels data in different classes should be well separated.

To quantify the separation between spectral classes only distance between means is insufficient since overlap will also be influenced by the standard deviations of the distributions. Therefore, a combination of both the distance between means and a measure of standard deviation is required (Richards, 1993). The mean controls the location of the distribution and the variance controls the spread of the data. When more than one feature is involved, then the multivariate normal distribution has to be used. In multivariate normal distribution instead of a single mean controlling the location of the distribution there is one mean for each feature making up a mean vector. The multivariate equivalent of the variance – covariance matrix which represents the variability of pixel values for each feature within a particular class and the correlations between the features. These two parameters are used to describe each class and computed for each sample. There are many class separability measures among them Bhattacharyya distance and Jeffries-Matusita distance is used in this study.

Bhattacharyya distance measures the similarity of two probability distributions and it is used to determine the separability of classes in classification. The Bhattacharyya distance of multispectral band is expressed as:

$$B_{\alpha\beta}^{(y)} = \frac{1}{8} (\mu_{\alpha} - \mu_{\beta})^t \left(\frac{C_{\alpha} + C_{\beta}}{2} \right)^{-1} (\mu_{\alpha} - \mu_{\beta}) + \frac{1}{2} \ln \frac{\left| \frac{C_{\alpha} + C_{\beta}}{2} \right|}{\sqrt{|C_{\alpha}| \times |C_{\beta}|}} \quad (3.3)$$

In a similar way, the Bhattacharyya distance of panchromatic band is expressed with mean and variance as:

$$B_{\alpha\beta}^{(z)} = \frac{1}{8} (\mu_{\alpha} - \mu_{\beta})^t \left(\frac{\sigma_{\alpha}^2 + \sigma_{\beta}^2}{2} \right)^{-1} (\mu_{\alpha} - \mu_{\beta}) + \frac{1}{2} \ln \frac{\left| \frac{\sigma_{\alpha}^2 + \sigma_{\beta}^2}{2} \right|}{\sqrt{|\sigma_{\alpha}^2| \times |\sigma_{\beta}^2|}}$$

$$B_{\alpha\beta}^{(z)} = \frac{(\mu_{\alpha} - \mu_{\beta})^2}{4(\sigma_{\alpha}^2 + \sigma_{\beta}^2)} + \frac{1}{2} \ln \left[\frac{\sigma_{\alpha}^2 + \sigma_{\beta}^2}{2\sigma_{\alpha}\sigma_{\beta}} \right] \quad (3.4)$$

When the number of classes is larger than two, an average Bhattacharyya distance is used

$$B_{avg}^{(y)} = \frac{2}{M(M-1)} \sum_{\alpha=1}^{M-1} \sum_{\beta=\alpha+1}^M B_{\alpha\beta}^{(y)} \quad (3.5)$$

$$B_{avg}^{(z)} = \frac{2}{M(M-1)} \sum_{\alpha=1}^{M-1} \sum_{\beta=\alpha+1}^M B_{\alpha\beta}^{(z)} \quad (3.6)$$

Jeffries-Matusita distance (JM) is introduced to transform the Bhattacharyya distance values to a specific range and the JM distance tends to overemphasising low separability values while suppress high separability values.

$$JM_{\alpha\beta}^{(y)} = 2 \left(1 - e^{-B_{\alpha\beta}^{(y)}} \right) \quad (3.7)$$

$$JM_{\alpha\beta}^{(z)} = 2 \left(1 - e^{-B_{\alpha\beta}^{(z)}} \right) \quad (3.8)$$

When the number of classes is larger than two, an average Jeffries-Matusita distance is used

$$JM_{avg}^{(y)} = \frac{2}{M(M-1)} \sum_{\alpha=1}^{M-1} \sum_{\beta=\alpha+1}^M JM_{\alpha\beta}^{(y)} \quad (3.9)$$

$$JM_{avg}^{(z)} = \frac{2}{M(M-1)} \sum_{\alpha=1}^{M-1} \sum_{\beta=\alpha+1}^M JM_{\alpha\beta}^{(z)} \quad (3.10)$$

Bhattacharyya distance values vary from 0 to ∞ where as Jeffries-Matusita distance vary from 0 to 2. If the mean and covariance of the two classes are the same, then $B_{\alpha\beta} = JM_{\alpha\beta} = 0$, this indicate that, it is

impossible to distinguish between the two classes based on spectral information only. On the contrary, when $B_{\alpha\beta} = \infty$ and $JM_{\alpha\beta} = 2$, indicating that the two classes are totally separated in feature space. Separability between the classes increases with increasing JM values and because of the saturating behaviour Jeffries-Matusita distances is preferred over the Bhattacharyya distance.

To see the effect of class separability of the multispectral image on the optimal λ and λ_p parameters different $JM^{(y)}$ values are chosen as 0.5, 1.0 and 1.9. An example of resulting class separability values between the classes for $JM^{(y)}=0.5$ and $JM^{(z)}=0.02$ is presented in table 3.1 and table 3.2 respectively. Similarly, to see the effects of class separability of the panchromatic image on the optimal λ and λ_p parameters the class separability of the panchromatic image is varying with $JM^{(z)}$ value of 0.5, 1.0 and 1.9.

Table 3.1: Jeffries-Matusita distance between the classes for $JM^{(y)} = 0.5$

	class1	class2	class3
class1	0	2	2
class2	2	0	0.5
class3	2	0.5	0

Table 3.2: Jeffries-Matusita distance between the classes for $JM^{(z)} = 0.02$

	class1	class2	class3
class1	0	1.811	1.914
class2	1.811	0	0.02
class3	1.914	0.02	0

Table 3.3: Relation between minimal and average Jeffries-Matusita distance in image y and z

$JM^{(y)}$	$JM^{(z)}$	$JM_{avg}^{(y)}$	$JM_{avg}^{(z)}$
0.5	0.02	1.5	1.24
1	0.02	1.67	1.25
1.9	0.02	1.97	1.26
2	0.5	2	1.43
2	1	2	1.6
2	1.9	2	1.9

It is possible to systematically control the class separability between spectral classes of the multispectral and panchromatic image, by expressing their class separability with the class parameter of the image x .

Class separability in the image x can be expressed with Bhattacharyya distance as:

$$B_{\alpha\beta}^{(x)} = \frac{1}{8} \left(\mu_{\alpha}^{(x)} - \mu_{\beta}^{(x)} \right)^2 \left(\frac{C_{\alpha}^{(x)} + C_{\beta}^{(x)}}{2} \right)^{-1} \left(\mu_{\alpha}^{(x)} - \mu_{\beta}^{(x)} \right) + \frac{1}{2} \ln \frac{\left| \frac{C_{\alpha}^{(x)} + C_{\beta}^{(x)}}{2} \right|}{\sqrt{|C_{\alpha}^{(x)}| |C_{\beta}^{(x)}|}} \quad (3.11)$$

To express the Bhattacharyya distance of image y in terms of the mean and covariance of image x first the relationship between the class parameters of the two images should be determined.

Assume image x and image y are spectrally similar. Then $\mu_{n,l}^{(y)} = \mu_{n,l}^{(x)}$

Where $\mu_{n,l}^{(y)}$ represents mean of image y and $\mu_{n,l}^{(x)}$ represents mean of image x .

$C_{n,l,m}^{(y)} = \frac{1}{S^2} C_{n,l,m}^{(x)}$ Where $C_{n,l,m}^{(y)}$ represent covariance of image y , $C_{n,l,m}^{(x)}$ represent covariance of image x and S is scale factor.

Here n represents the number of class, l and m represents the number of bands.

Then the Bhattacharyya distance of the Image y is expressed as:

$$B_{\alpha\beta}^{(y)} = \frac{1}{8} (\mu_{\alpha}^{(x)} - \mu_{\beta}^{(x)})^T \left(\frac{C_{\alpha}^{(x)} + C_{\beta}^{(x)}}{2S^2} \right)^{-1} (\mu_{\alpha}^{(x)} - \mu_{\beta}^{(x)}) + \frac{1}{2} \ln \frac{|C_{\alpha}^{(x)} + C_{\beta}^{(x)}|}{2\sqrt{|C_{\alpha}^{(x)}| |C_{\beta}^{(x)}|}} \quad (3.12)$$

Where α and β are the two spectral classes.

Similarly, to express the Bhattacharyya distance of image z in terms of mean and covariance of image x first the relationship between the class parameters of the two images should be determined.

It is assumed that, the spatial resolution of image x and image z are the same.

$\mu_k^{(z)} = \frac{1}{N_b} \sum_{l=1}^{N_b} \mu_{k,l}^{(x)}$, where $\mu_k^{(z)}$ represents mean of the panchromatic image z and

N_b Represents number bands.

$C_k^{(z)} = \frac{1}{N_b^2} \sum_{l=1}^{N_b} \sum_{m=1}^{N_b} c_{k,l,m}^{(x)}$, where $C_k^{(z)}$ represents covariance of image z .

Then the Bhattacharyya distance of the image z is expressed as:

$$B_{\alpha\beta}^{(z)} = \frac{1}{8} \left(\frac{1}{N_b} \sum_{l=1}^{N_b} (\mu_{\alpha}^{(x)} - \mu_{\beta}^{(x)}) \right) \left(\frac{\frac{1}{N_b^2} \sum_{l=1}^{N_b} \sum_{m=1}^{N_b} (C_{\alpha}^{(x)} + C_{\beta}^{(x)})}{2} \right)^{-1} \left(\frac{1}{N_b} \sum_{l=1}^{N_b} (\mu_{\alpha}^{(x)} - \mu_{\beta}^{(x)}) \right) + \frac{1}{2} \ln \frac{\frac{1}{N_b^2} \sum_{l=1}^{N_b} \sum_{m=1}^{N_b} |C_{\alpha}^{(x)} + C_{\beta}^{(x)}|}{2\sqrt{\frac{1}{N_b^2} \sum_{l=1}^{N_b} \sum_{m=1}^{N_b} |C_{\alpha}^{(x)}| \frac{1}{N_b^2} \sum_{l=1}^{N_b} \sum_{m=1}^{N_b} |C_{\beta}^{(x)}|}} \quad (3.13)$$

Equation 3.12 and 3.13 is used to control the class separability values in the multispectral and panchromatic image only by changing mean and covariance of the image x in the synthetic image generation.

3.3. MRF

The theory of MRF including its formulation described below is adapted from (Tolpekin et al., 2010) with some minor changes.

The advantages of using MRF models is that, the prior information and the observed data from different sources can be integrated through the use of suitable energy function. The super resolution map (SR) map c that corresponds to the maximum a posteriori probability $p(c | y, z)$ solution for c given observed data y and z can be computed with Bayes theorem from prior probability $p(c)$ and conditional probabilities $p(y | c)$ and $p(z | c)$ as:

$$p(c | y, z) \propto p(c)p(y | c)p(z | c) \quad (3.14)$$

Assume y and z are conditionally independent given c and due to the equivalence of MRF and Gibbs random field, the probabilities can be specified by means of energy functions as:

$$p(c) = \frac{1}{A_1} \exp\left(\frac{-U(c)}{T}\right) \quad (3.15)$$

$$p(y | c) = \frac{1}{A_2} \exp\left(\frac{-U(y | c)}{T}\right) \quad (3.16)$$

$$p(z | c) = \frac{1}{A_3} \exp\left(\frac{-U(z | c)}{T}\right) \quad (3.17)$$

$$p(c | y, z) = \frac{1}{A_4} \exp\left(\frac{-U(c | y, z)}{T}\right) \quad (3.18)$$

Where A_1 , A_2 , A_3 and A_4 are normalizing constants, T is a constant termed the temperature, $U(c)$, $U(y | c)$, $U(z | c)$ and $U(c | y, z)$ are the prior, two conditional and the posterior energy functions respectively. The resulting expression, when rewriting the Bayes formula for energy function is

$$U(c | y, z) = \lambda U(c) + (1 - \lambda)(\lambda_p U(z | c) + (1 - \lambda_p)U(y | c)) \quad (3.19)$$

Where $0 \leq \lambda < 1$, is a parameter which balances the contribution of prior and conditional energy functions. $\lambda = 0$ indicates that the contextual information is ignored in the classification and only the conditional energy is used. Whereas, $\lambda = 0.5$ indicates that equal weights are assigned to the prior and conditional energy. $\lambda = 1$ indicates, the conditional energy is ignored and only contextual information is used in the classification which results a similar class.

Likewise, $0 \leq \lambda_p \leq 1$, is a parameter which balances the contribution of the panchromatic and multispectral conditional energy functions. $\lambda_p = 0$ indicates the conditional energy from the panchromatic image is ignored and the classification is done only by using the likelihood energy from the multispectral band. Whereas, $\lambda_p = 1$ indicates that, the conditional energy from the multispectral image is ignored and only the conditional energy from the panchromatic image is used. This is similar to maximum likelihood classification of the panchromatic band.

3.3.1. Prior energy function

The prior energy is modelled as the sum of pair-site interactions. (Li, 2001)

$$U(C) = \sum_{i,j} U(c(a_{j|i})) = \sum_{i,j} \sum_{l \in N(a_{j|i})} w(a_l) I(c(a_{j|i}), c(a_l)) \quad (3.20)$$

Here, $N(a_{j|i})$ is the neighbourhood system, $U(c(a_{j|i}))$ is the local contribution to the prior energy from the fine resolution pixel $c(a_{j|i})$, $w(a_l)$ represents the weight of the contribution from pixel $a_l \in N(a_{j|i})$ to the prior energy and $I(c_1, c_2)$ takes the value 0 if $c_1 = c_2$ and 1 otherwise.

This prior model penalizes the occurrence of pixels with different class labels in the neighbourhood system N and the weights $w(a_l)$ are chosen inversely proportional to the distance $d(a_l)$ between the central pixel $a_{j|i}$ and the pixel a_l as

$$w(a_l) \propto \frac{1}{d(a_{j|i}, a_l)} \quad (3.21)$$

3.3.2. Conditional energy functions

The likelihood energies consider the closeness of observed pixel values to each land cover class. The spectral values x of a class α is modelled with the normal distribution and it is assumed that the spectral values $x(a_{ji})$ of the S^2 fine resolution pixel a_{ji} are independent: spatially uncorrelated and identically distributed given their class association.

The conditional energy is defined using the mean vector μ_α of class α and the covariance matrix as:

$$U(y | c) = \sum_{i,j} U(y(d_i) | c(a_{ji}))$$

Where $U(y(d_i) | c(a_{ji}))$ is the local contribution to the likelihood energy from the fine resolution pixel $c(a_{ji})$

$$U(y | c) = \sum_{i,j} \left[\frac{1}{2} (y(d_i) - \mu_i)^t C_i^{-1} (y(d_i) - \mu_i) + \frac{1}{2} \ln |C_i| \right] \quad (3.22)$$

The values μ_i and C_i are modelled as a linear combination of mean vectors and covariance matrices based on proportions of respective land cover classes inside the coarser resolution pixel.

Similarly, the panchromatic conditional energy $U(z | c)$ is defined as:

$$U(z | c) = \sum_{i,j} \frac{1}{2} \left[\frac{(z(a_{ji}) - \mu_\alpha)^2}{\sigma_\alpha^2} + \ln \sigma_\alpha^2 \right] \quad (3.23)$$

Where $z(a_{ji})$ is the pixel value, μ_α is mean of the class α and σ_α is the standard deviation of the class α .

3.4. Estimation of λ and λ_p parameter

λ and λ_p are internal parameters of the MRF based SRM and it needs to be properly estimated before the method is applied to obtain higher classification accuracy. Based on the result obtained in Tolpekin, et al.(2010), it is assumed that, the incorporation of information from the panchromatic band improves the classification results when compared to classification with out information from the panchromatic band, if it is properly estimated. λ_p parameter is introduced to balance the contribution of likelihood energy from the multispectral and panchromatic image. MRF parameter estimation for SRM is influenced by class separability and scale factor information(Tolpekin and Stein, 2009). In this study models are developed based on local energy balance analysis and by relating the optimal λ and λ_p parameter estimation to the class separability and scale factor information.

First assume correct classification of classes and consider a case when a fine-resolution pixel a_{ji} with a true class label $c(a_{ji}) = \alpha$, to which is assigned a different class label β and all the other pixels are selected correctly. Then the resulting change in likelihood energy in multispectral and panchromatic band is as follows:

The change in likelihood energy of the panchromatic band is:

$$\Delta U^{(z)} = U^{(z)}(c = \alpha) - U^{(z)}(c = \beta) \quad (3.24)$$

Where c is the label of the fine resolution pixel, α and β are spectral classes.

The panchromatic energy with the true class label $c(a_{ji}) = \alpha$ is expressed as:

$$U^{(z)}(c = \alpha) = \frac{1}{2} \left[\frac{(\mu_\alpha - \mu_\alpha)^2}{\sigma_\alpha^2} + \ln \sigma_\alpha^2 \right] \quad (3.25)$$

Where σ represents standard deviations

The panchromatic energy with the wrong class label $c(a_{ji}) = \beta$ is expressed as:

$$U^{(z)}(c = \beta) = \frac{1}{2} \left[\frac{(\mu_\alpha - \mu_\beta)^2}{\sigma_\beta^2} + \ln \sigma_\beta^2 \right] \quad (3.26)$$

The change in energy when the fine resolution pixel label $c(a_{ji})$ changes from α to β is expressed as:

$$\Delta U_{\alpha\beta}^{(z)} = \frac{1}{2} \left[\ln \sigma_\alpha^2 - \ln \sigma_\beta^2 \right] - \frac{(\mu_\alpha - \mu_\beta)^2}{2\sigma_\beta^2} = \ln \frac{\sigma_\alpha}{\sigma_\beta} - \frac{(\mu_\alpha - \mu_\beta)^2}{2\sigma_\beta^2} \quad (3.27)$$

To simplify the modelling, it is assumed that, the variances of the two classes are the same.

$\sigma_\alpha = \sigma_\beta$, then the above expression is simplified as:

$$\Delta U_{\alpha\beta}^{(z)} = \frac{(\Delta\mu)^2}{2\sigma^2} \quad (3.28)$$

Where $\Delta\mu$ is the difference between the mean of the two spectral classes.

Similarly, equation 2.8 is simplified as:

$$B_{\alpha\beta}^{(z)} = \frac{(\Delta\mu)^2}{8\sigma^2} \quad (3.29)$$

Then, the change in energy of the panchromatic band is expressed with the Bhattacharyya distance as:

$$\Delta U_{\alpha\beta}^{(z)} = 4B_{\alpha\beta}^{(z)} \quad (3.30)$$

The change of the fine resolution pixel label $c(a_{ji})$ from α to β also causes the change of the composition of the coarse-resolution pixel d_i and thus, to a change in the mean vector and the covariance matrix for this pixel value.

The change in likelihood energy of the multispectral band is calculated with mean and covariance as:

$$\Delta U_{\alpha\beta}^{(y)} = \frac{1}{2} \left(\frac{\mu_\beta - \mu_\alpha}{S^2} \right)^t \left(\frac{C_\alpha}{S^2} \right)^{-1} \left(\frac{\mu_\beta - \mu_\alpha}{S^2} \right) \quad (3.31)$$

$$\Delta U_{\alpha\beta}^{(y)} = \frac{1}{2S^2} (\Delta\mu)^t C^{-1} (\Delta\mu) \quad (3.32)$$

With the same assumption for the classes α and β , Equation (3.3) of the Bhattacharyya distance between classes in the multispectral band is simplified as:

$$B_{\alpha\beta}^{(y)} = \frac{1}{8} (\Delta\mu)^t C^{-1} (\Delta\mu) \quad (3.33)$$

Then, the change in energy of the multispectral band is expressed with the Bhattacharyya distance as:

$$\Delta U_{\alpha\beta}^{(y)} = \frac{4B_{\alpha\beta}^{(y)}}{S^2} \quad (3.34)$$

Subsequently, the value of λ_p is determined from the balance between the change in likelihood energy of the panchromatic and multispectral band with:

$$\lambda_p \Delta U_{\alpha\beta}^{(z)} = (1 - \lambda_p) \Delta U_{\alpha\beta}^{(y)} \quad (3.35)$$

$$\lambda_p = \frac{1}{1 + \frac{S^2 \times B_{\alpha\beta}^{(z)}}{B_{\alpha\beta}^{(y)}}} \quad (3.36)$$

Likewise, to estimate the smoothness parameter λ , consider also a case when a fine-resolution pixel a_{ji} with a true class value $c(a_{ji}) = \alpha$, to which is assigned a different class label β . Then the change of the local prior energy when the pixel label $c(a_{ji})$ changes from α to β is

$$\Delta U_{\alpha\beta}^p = q \left| \sum_{l \in N(a_{ji})} \phi(a_l) [I(\beta, c(a_l)) - I(\alpha, c(a_l))] \right| = q\gamma \quad (3.37)$$

The value of γ depends on the neighbourhood system size, the configuration of class labels $c(a_l)$ in the neighbourhood $N(a_{ji})$ and the choice of power-law index n . (Tolpekin and Stein, 2009).

The total change in likelihood energy when the pixel label $c(a_{ji})$ changes from α to β will be

$$\begin{aligned} \Delta U_{\alpha\beta}^{tl} &= \Delta U_{\alpha\beta}^{(z)} + \Delta U_{\alpha\beta}^{(y)} \\ \Delta U_{\alpha\beta}^{tl} &= 4B_{\alpha\beta}^{(z)} + \frac{4 \times B_{\alpha\beta}^{(y)}}{S^2} \end{aligned} \quad (3.38)$$

Then, the smoothness parameter λ can be determined from the balance between the change in the prior and total likelihood energy values.

$$\begin{aligned} \Delta U_{\alpha\beta}^{tl} &= 4B_{\alpha\beta}^{(z)} + \frac{4 \times B_{\alpha\beta}^{(y)}}{S^2} = q\gamma \\ \Delta U_{\alpha\beta}^{tl} &= 4B_{\alpha\beta}^{(z)} + \frac{4 \times B_{\alpha\beta}^{(y)}}{S^2} = \left(\frac{\lambda}{1-\lambda} \right) \gamma \end{aligned} \quad (3.39)$$

Finally, solving for λ

$$\lambda = \frac{1}{1 + \frac{S^2 \times \gamma}{4B_{\alpha\beta}^{(y)} + 4S^2 B_{\alpha\beta}^{(z)}}} \quad (3.40)$$

If the class separability in the multispectral and panchromatic images is small, then the smoothness parameter λ value will also be small. On the contrary if the class separability in the multispectral and panchromatic image is large, then the smoothness parameter λ value will be high.

3.5. Simulated Annealing

Once the global energy constructed and the optimal parameters have been determined, the next step of SRM is to find the minimum of the total energy in map estimate. To achieve this simulated annealing algorithm (SA) is employed to find the map solution. SA is a stochastic algorithm used to find a good optimization problem based on the use of random numbers and probability statistics. Temperature and σ are the main parameters in the simulated annealing algorithm that control the process of optimization

and mostly these parameters are estimated experimentally. In this study a power-law annealing schedule is used, where the temperature at the iteration n is changed according to

$$T_n = \sigma \times T_{n-1} \quad (3.41)$$

The parameter σ controls the rate of temperature decrease and the parameter T controls the randomness of the optimization algorithm. High temperature means high randomness and low temperature means less randomness and the high temperature increases the probability of a pixel label to be replaced with a new label. Pixel values are updated with the Metropolis-Hastings sampler then the number of successful updates is counted, being the number of updates that lead to a change of pixel value. If the number of successful updates is below 0.1% of the total number of pixels during three consecutive iterations the optimization is stopped.

The initial super resolution map is obtained by first applying linear spectral unmixing, to estimate the proportion of each class within each coarse resolution pixel d_i . Then the coarse resolution pixel is filled with randomly distributed fine resolution pixels of each class according to their proportions. Finally, by optimizing the spatial dependence the SRM is generated. In SRM, the optimal values of the parameters T_0 and σ in the annealing schedule depends on the scale factor and class separability. (Tolpekin and Stein, 2009)

3.6. Accuracy assessment and performance analysis.

It is necessary to evaluate the accuracy of the classification result against the reference data to assure its fitness for use of the intended application. In this context the accuracy is defined as the level of agreement between labels assigned by the SRM with the reference data. The common way of representing the classification accuracy is with the form of an error matrix which compares the classification results to the corresponding reference data on class by class basis. (Richards, 1993).

The kappa coefficient is one of the quality measures that is derived from the whole error matrix information and estimates the overall agreement between the classified image and the reference data. The highest value of the kappa coefficient is 1 that means there is a perfect agreement between the classified image and the reference image where as kappa value 0 shows the classified image is completely different from the reference image. In this thesis the accuracy of the results are measured in kappa coefficient and the performance of the models are checked with the numerical optimal values obtained from the experiment. To do this, experiments are conducted to identify the optimal λ and λ_p parameter for each scale factors and class separability. An optimal λ and λ_p values are those values that results the highest kappa value. After that, optimal λ and λ_p range is specified. Finally the predicted values are compared to the experimentally determined optimal values.

4. RESULTS

This chapter presents the optimal λ and λ_p parameter estimation results obtained in MRF based super resolution mapping and the analysis of the results. Section 4.1 presents the experimental results conducted to check how the scale factor and class separability value affects the optimal parameters of λ and λ_p . In section 4.2 the results obtained from the models are presented and analyzed. Section 4.3 summarizes the main findings of the results.

4.1. Experimental results from synthetic datasets

To understand how the optimal λ and λ_p parameters are affected with different scale and class separability values, several synthetic images were generated as mentioned in Section 3.2. This process of image generation simplifies the complexity of real image. Proper parameter identification and estimation are the most crucial part of this study and for this purpose, several experiments have been done and the quality of the result is measured using kappa coefficient. In each process the algorithm is run a minimum of ten times to get a consistent result and the mean kappa value has been determined.

4.1.1. The effect of scale factor on the optimal λ and λ_p parameter.

The intension of doing this experiment is to observe the effects of scale factor on the optimal parameter of λ and λ_p . This experiment is carried out using scale factor values of 2, 4,5,6,8, 10 and combination of λ values which range from 0.1-0.9 and pan value from 0.0-0.9 are used. The smoothness parameter λ is used as a balancing factor between the prior and likelihood energy function whereas λ_p is used as a balancing factor between the multispectral and panchromatic likelihood energy functions. The first experiments are conducted to find the optimal λ and λ_p combinations, that give maximum classification accuracy and to quantify there relation with the scale factor. Figure 4.1 (b) shows an example of the super resolution map result obtained with the optimal combination of λ and λ_p at scale factor 2.

Figure 4.2, shows that, as the scale factor increases the optimal smoothness parameter λ value decreases. This is mainly due to the no of sub-pixels in the coarser resolution pixels increases with scale factor. This makes the coarser resolution pixels highly mixed and hence, using larger smoothness parameter decreases the classification accuracy. However, for smaller scale factors, using more contextual information increases the classification accuracy.

As can be seen from figure 4.2, at the lower scale factor, λ_p is very important to obtain higher classification accuracy and almost equal amount of likelihood energy is contributed from the panchromatic and multispectral image. However, as the scale factor increases the optimal λ_p value decreases. This means that, the contribution of the likelihood energy from the panchromatic image decreases. In other words, for larger scale factors the inclusion of more likelihood energy from the panchromatic band decreases the quality of SRM.

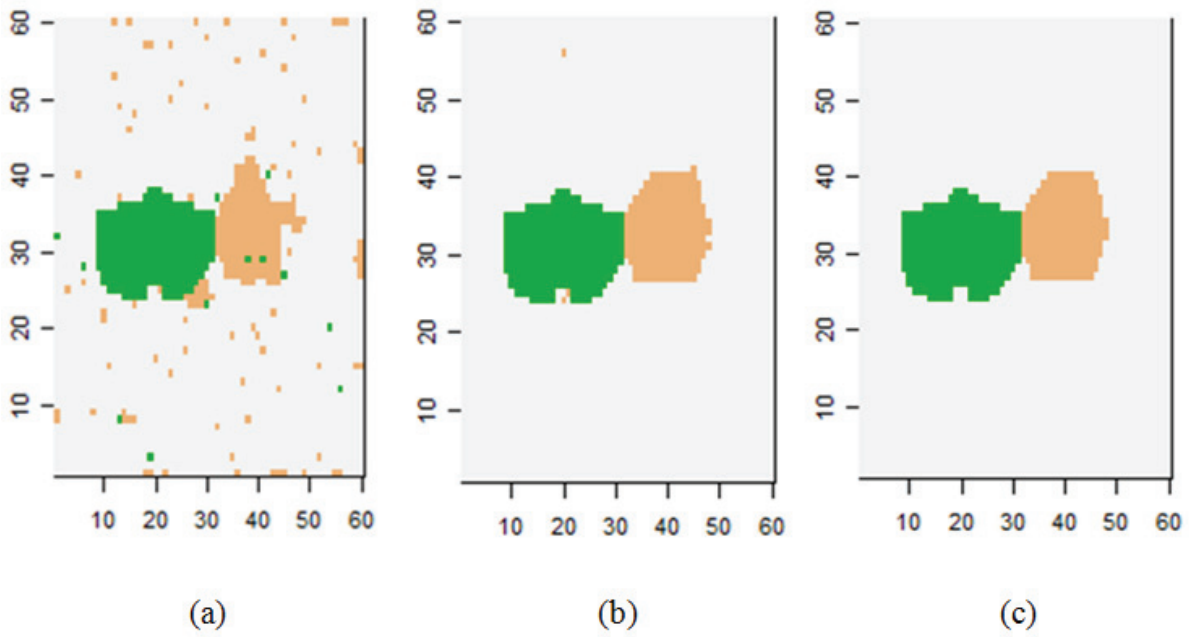


Figure 4.1: super resolution map result. (a) lowest kappa obtained, (b) highest kappa obtained, (c) reference image.

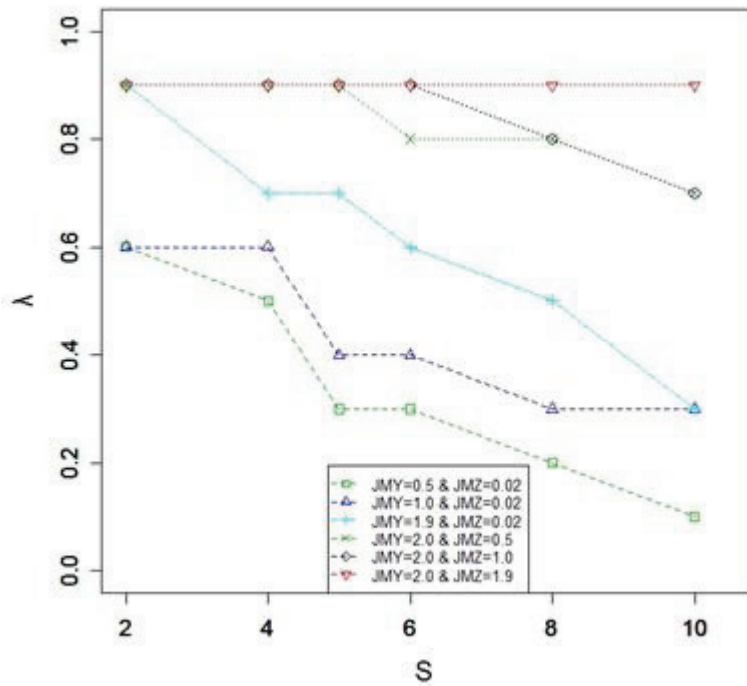


Figure 4.2: Optimal λ value at different scale and class separability in the multispectral and panchromatic image. Lines are added to facilitate the interpretation of the data.

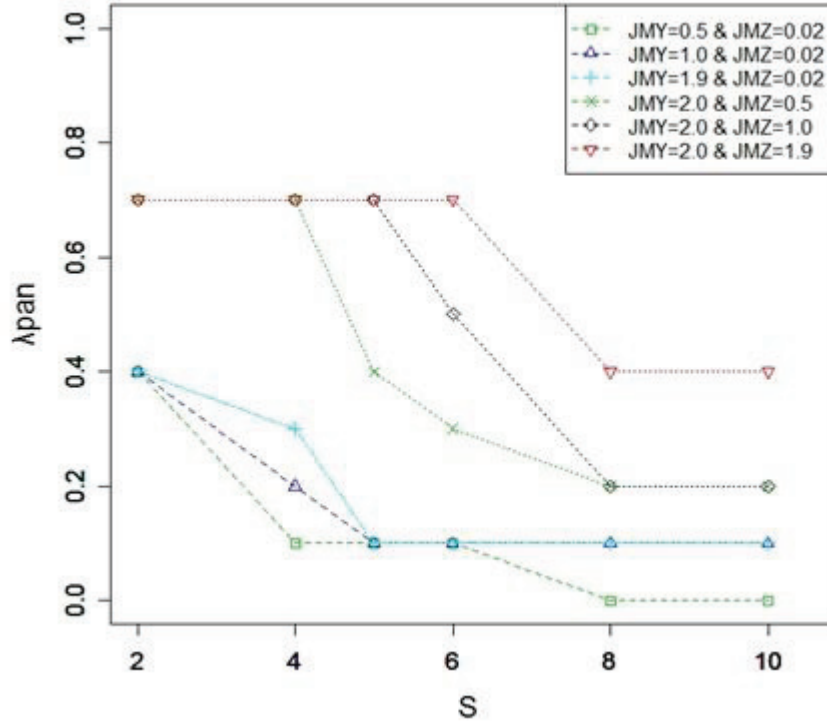


Figure 4.3: optimal λ_{pan} value at different scale and class separability in the multispectral and panchromatic image. Lines are added to facilitate the interpretation of the data.

4.1.2. The effect of class separability on the optimal λ and λ_{pan} parameter.

This experiment is performed to see the effects of class separability between spectral classes in the multispectral and panchromatic image on the optimal values of λ and λ_p parameters in MRF based super resolution mapping. This is carried out first by fixing the class separability values in the panchromatic image while varying class separability values in the multispectral image. To see the effects, the synthetic images is constructed with class separability values in the multispectral image of $JM^{(y)} = 0.5, 1.0$ and 1.9 with fixed class separability value of $JM^{(z)} = 0.02$ in the panchromatic image, for different scale factor values. Similarly, to see the effects of the panchromatic image the class separability of the multispectral image is fixed while the class separability of the panchromatic image is varying with $JM^{(z)}$ value of $0.5, 1.0$ and 1.9 .

As can be seen from Figure 4.1, when class separability between spectral classes in the multispectral and panchromatic image is smaller, the optimal smoothness parameter λ value becomes less. This indicates that, the contribution of the likelihood energy from the multispectral and panchromatic image to the global energy should be higher to get better classification accuracy. In another way, when the class separability between spectral classes in the multispectral image increases the optimal smoothness parameter λ value also increases. This shows that, if the spectral classes are well separated using more contextual information further increases the classification accuracy. Moreover, when the class separability in the panchromatic image increases the smoothness parameter λ value also slightly increases. These results shows that, class separability and smoothness parameters have direct relationships. If there is

maximum class separability between spectral classes in the multispectral and panchromatic image the smoothness parameter λ value becomes higher.

In general, for most of the scale factors when the class separability in the multispectral band increases, the optimal λ value also increases. Nevertheless, increase in class separability of the panchromatic image does not significantly affect the optimal λ value. As observed from Figure 4.2, the increase in the class separability between spectral classes results in increase the optimal λ_p parameter. This indicates that, if spectral classes are well separated, using more likelihood energy from the panchromatic band increases the classification accuracy.

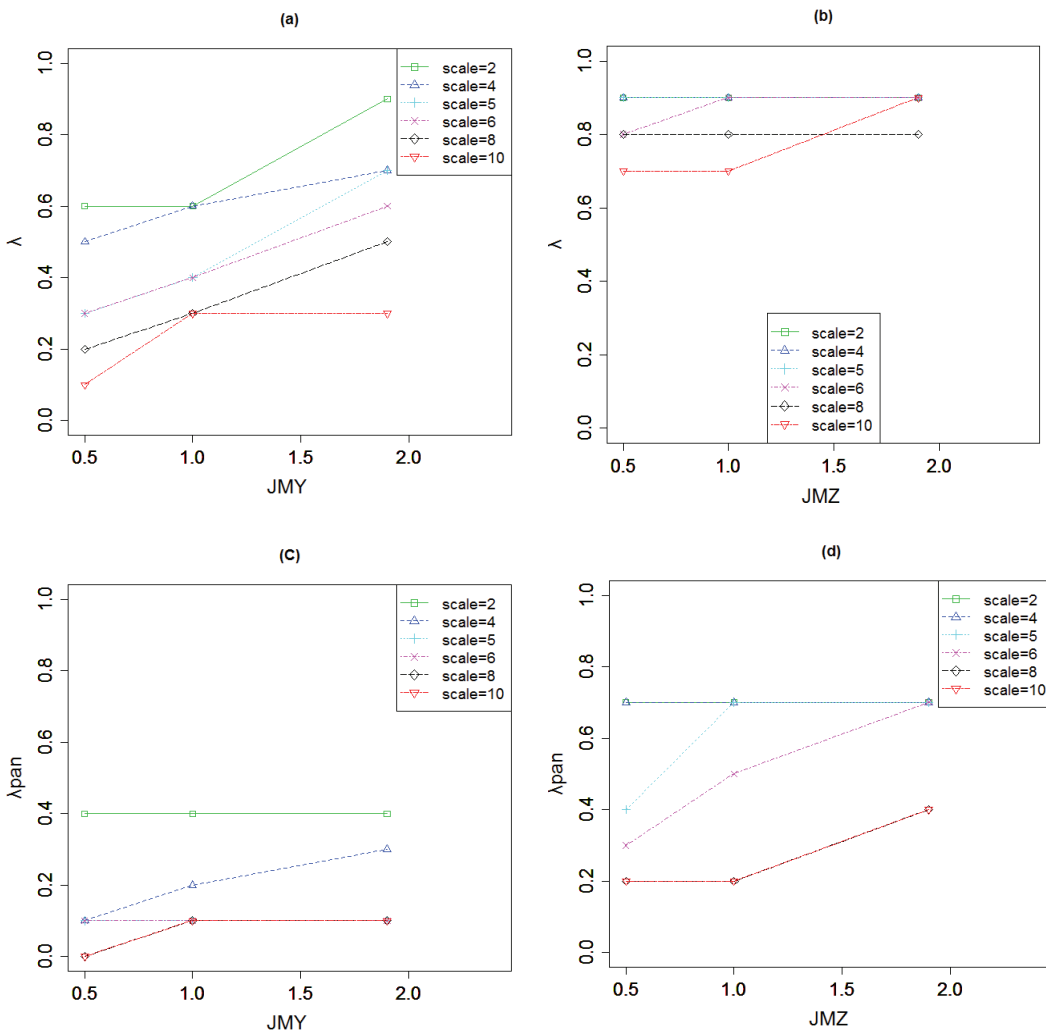


Figure 4.4: The effect of class separability on the optimal values of λ and λ_{pan} . (a) and (c) show the change of the optimal λ and λ_{pan} values respectively, when class separability in the multispectral image changes for a fixed JMZ value of 0.02. Here, (b) and (d) show the change of the optimal λ and λ_{pan} values respectively, when class separability in the panchromatic image changes for a fixed JMY value of 2.0.

It is clearly seen from Figure 4.3 (a), when the class separability between spectral classes in the multispectral image increases, the optimal smoothness parameter λ value as well increases significantly. This indicates good class separability between spectral classes make the prior model provide more information for the classification. The optimal smoothness parameter λ value is also slightly affected with the increase in class separability between spectral classes in the panchromatic image as shown in figure 4.3 (b). On the contrary, if class separability in the panchromatic image increases the optimal λ_p value increases more as shown in figure 4.3 (d). Where as, Figure 4.3 (c) show, optimal λ_p value is less affected with the increase in class separability in the multispectral image.

4.1.3. Effect of scale factor and class separability on the classification accuracy kappa

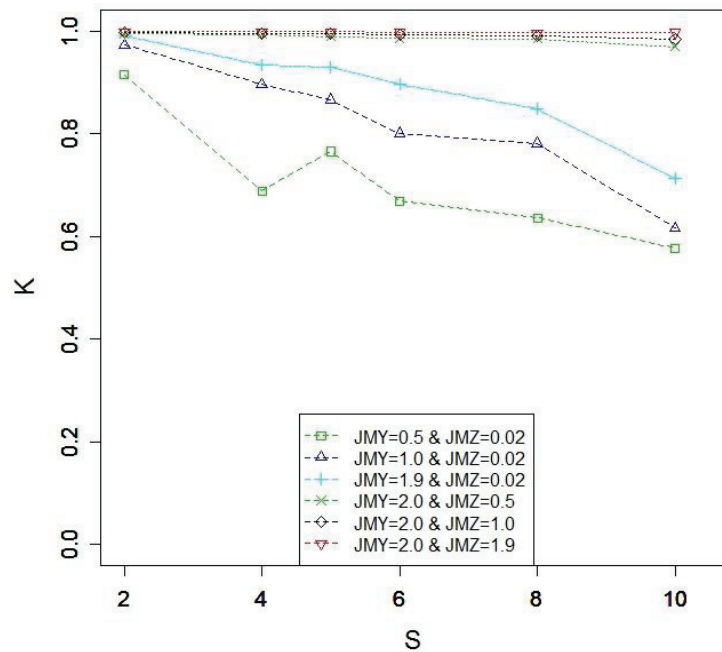


Figure 4.5: The effect of scale factor and class separability in the classification accuracy kappa value.

Figure 4.5 indicates that, classification accuracy kappa value show significant increase, when class separability between spectral classes in the multispectral image increases. On the contrary, very low class separability leads to confusion between spectral classes, which result in inferior classification accuracy. It is also noticed that, when class separability between spectral classes in the panchromatic image increases the classification accuracy kappa value increases to some extent. Generally it is observed from the figure that, when the scale factor increases the classification accuracy kappa value decreases. nevertheless, at larger class separability in multispectral and panchromatic image the classification accuracy kappa value does not show big variations, the value more or less remain the same for different scale factors. The effect of scale factor on the classification accuracy kappa value is very small for widely separable classes. However, it affects more, if the class separability between spectral classes in the multispectral and panchromatic image is small.

As it is observed from figures A.1, A.2 and A.3 in appendix A, the value of λ and λ_p has an effect on the value of kappa for each combination of S, $JM^{(y)}$ and $JM^{(z)}$. For every combination of S, $JM^{(y)}$ and $JM^{(z)}$, there exist an optimal λ and λ_p combination where the classification accuracy attains the maximal value K_{max} . There is also a range of λ and λ_p values that results classification accuracy kappa close to k_{max} . In this study the range for λ and λ_p values is defined in such away that the kappa value obtained with those λ and λ_p values is greater than $0.85k_{max}$.

4.2. λ and λ_p estimation results from synthetic image

In this section the results estimated with the models for optimal λ and λ_p parameters are presented. As it is observed from figure 4.5, 4.6 and 4.7, the prediction of λ and λ_p value with the model fits the empirical data obtained from the experiments.

To estimate the parameters with the model only scale factor and class separability information is required (see equation 3.22 and 3.26). The class separability between spectral classes in the multispectral and panchromatic image is calculated by using equation 2.7 and 2.8 respectively. In figure 4.5, the parameters are estimated by considering the minimum class separability between all classes and if class separability values between different pairs of classes vary strongly an average class separability value can be used. The estimation results based on the average class separability values of all classes are listed in table C.1 in appendix C. Some adjustment factors are used in the model developed in (3.22 & 3.26) to best fit the estimation with the experimental data. Identification of λ and λ_p is not very precise because the estimation was done by increasing λ and λ_p values with step sites of 0.1.

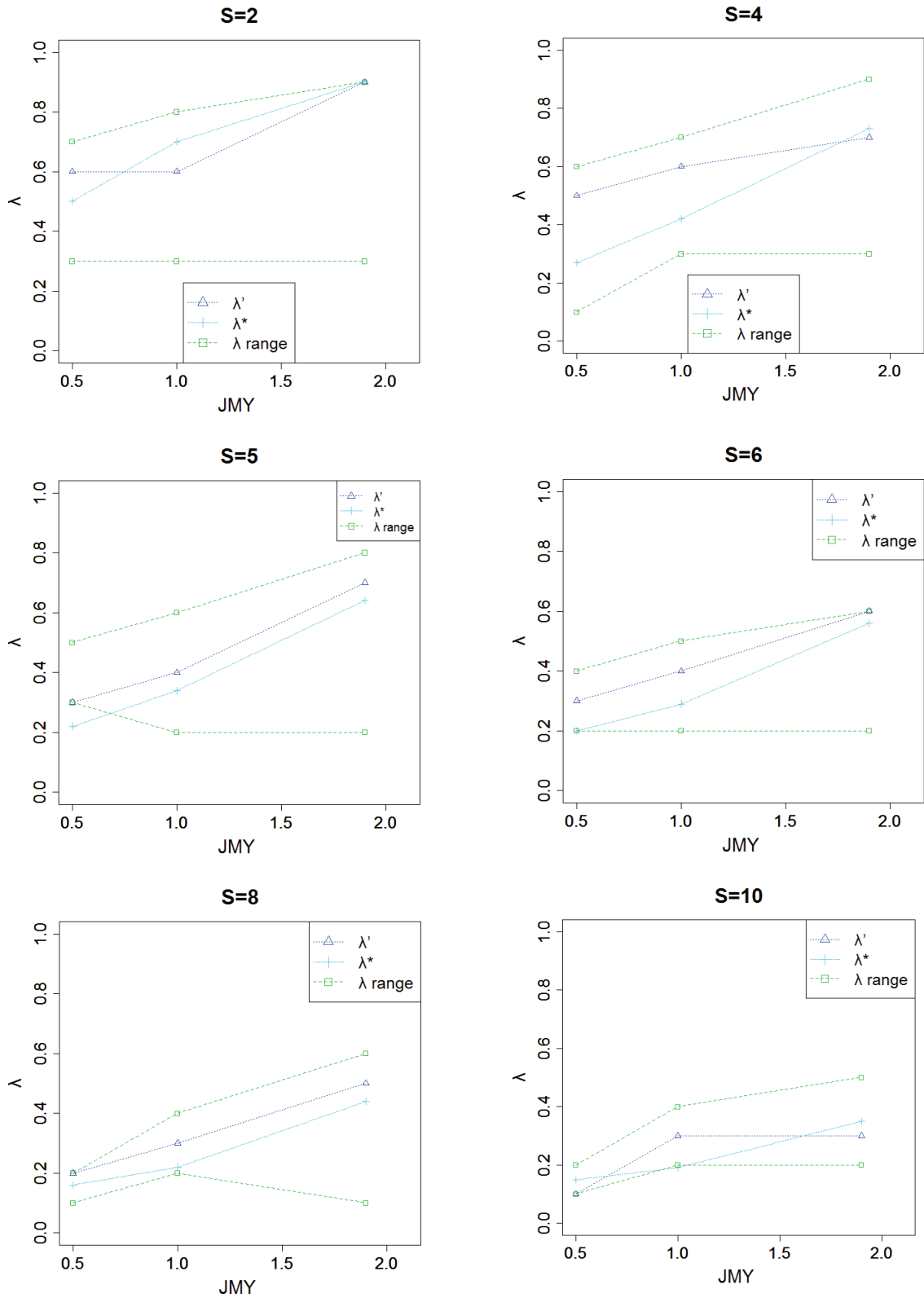


Figure 4.6: Optimal λ values for varying parameters S and JMY . λ' : The experimentally determined optimal values. λ range: The range of λ corresponding to $K \geq 0.85 k_{\max}$. λ^* : The estimation from the model.

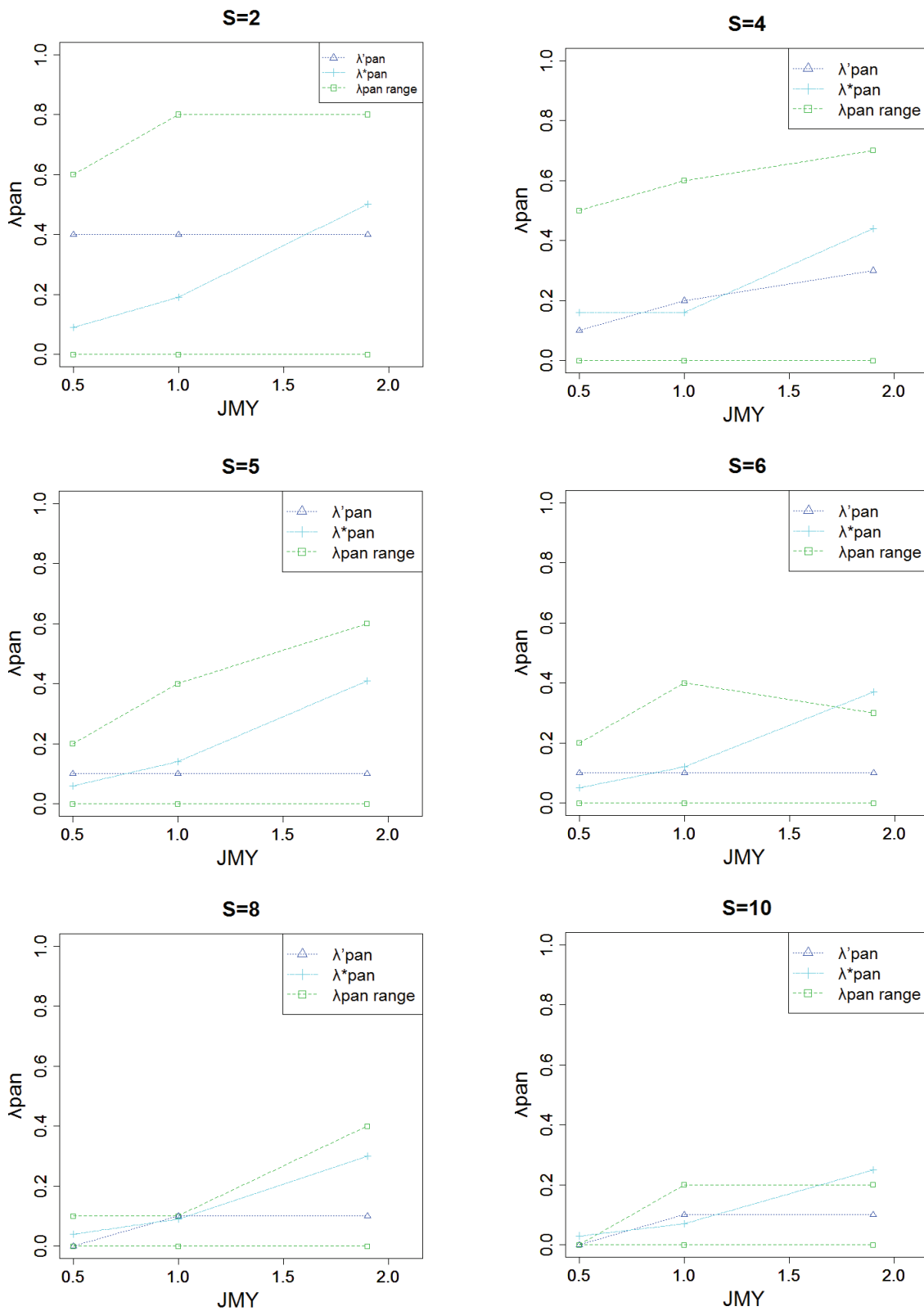


Figure 4.7: Optimal λ_{pan} values for varying parameters S and JMY. λ'_{pan} : The experimentally determined optimal values. λ_{pan} range: The range of λ_{pan} corresponding to $K \geq 0.85K_{max}$. λ^*_{pan} : The estimation from the model

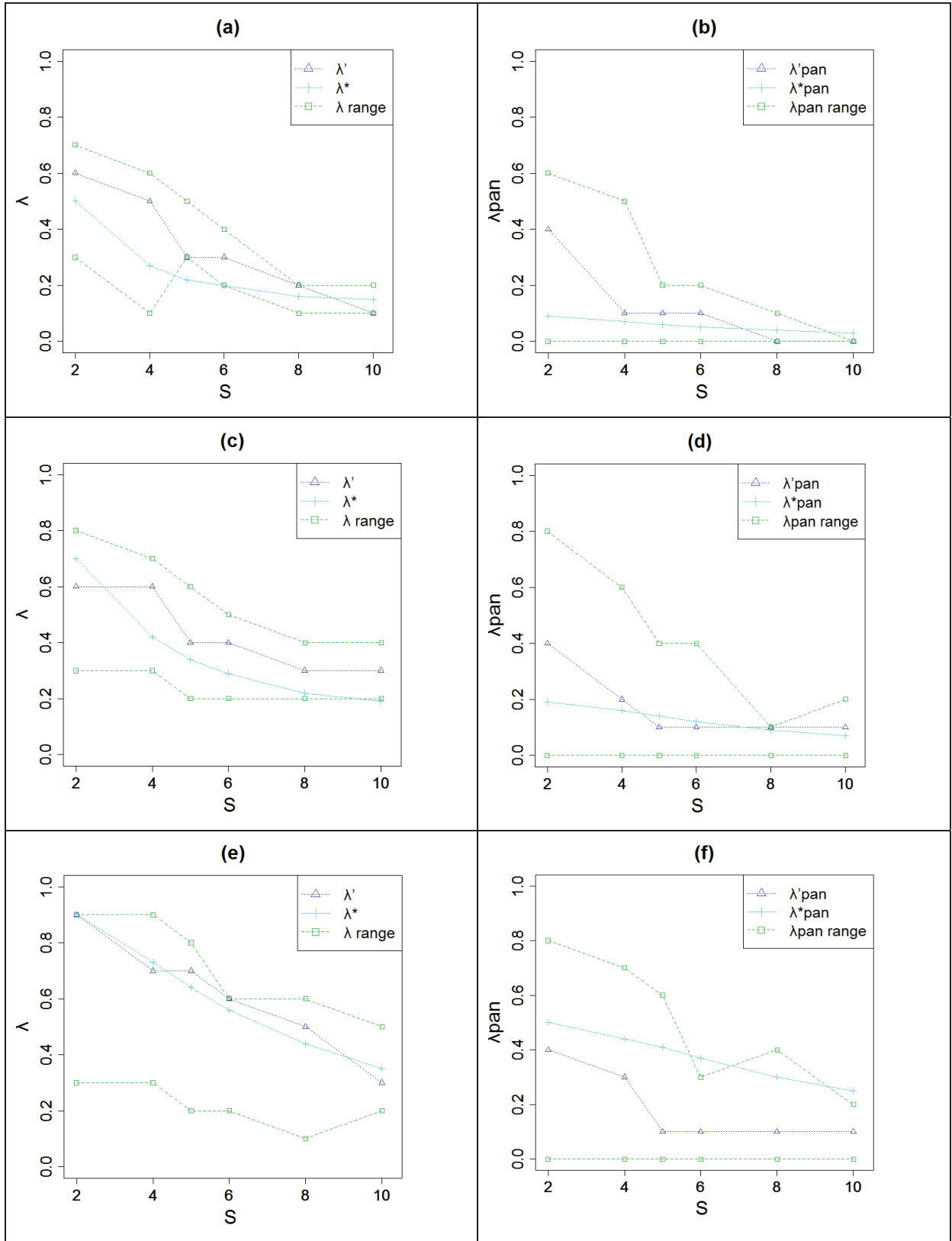


Figure 4.8: Optimal values for varying parameters S and JMY. λ'_{pan} & λ' : The experimentally determined optimal values. λ^*_{pan} & λ^* : The estimation from the model. $\lambda_{pan\ range}$ & λ_{range} : The range of λ_{pan} and corresponding to $k \geq 0.85k_{max}$. Here (a) and (b) for JMY value of 0.5 and (c) and (d) for JMY value of 1.0. (e) and (f) for JMY value of 1.9 for a fixed JMZ value of 0.02.

4.3. parameter estimation result from the real image

The applicability of the models for real image is tested with the result obtained from the experiment, conducted by Wijeratna Nanthamuni (2010 MSC student, who works on the title, Integration of Spatial and Spectral data of very high resolution imagery for building footprint detection using Super Resolution Mapping). The optimal λ and λ_p results obtained from the experiment are 0.8 and 0.3 respectively and the models estimates optimal λ value equal to 0.98 and for λ_p value 0.3. This shows that the models can estimate the optimal parameters for real images with in a tolerable range.

4.4. Summary of observation from the results

The result from the experiments proof that, the optimal λ and λ_p parameters have a direct relation ship with class-separability and have an inverse relation ship with the scale factor values. There exist an optimal combination of λ and λ_p values, where the classification accuracy attains the maximum value k_{\max} , for each SRM input parameters scale factor and class separability. The result obtained from the models developed in this research for the optimal λ and λ_p parameter estimation agrees with the empirical data. The likelihood information obtained from the panchromatic band increases the classification accuracy at the lower scale factors.

5. DISCUSSIONS, CONCLUSION AND RECOMMENDATION

5.1. Discussions

The main objective of this study is to develop models based on local energy balance analysis for the estimation of optimal λ and λ_p parameters. The smoothness parameter λ acts as a balancing factor between the prior and conditional energy in the MRF model. Whereas λ_p controls the balance between the likelihood energy from the panchromatic and multispectral band. In this study the effects of class separability and scale factor on the optimal parameters of λ and λ_p in the MRF based super resolution mapping are studied.

The obtained result from the experiments show that the optimal λ and λ_p parameters are influenced with scale factor and class separability values. As the scale factor increases the optimal smoothness parameter λ value decreases. This is mainly due to the number of sub-pixels in the coarser resolution pixels increases with scale factor and makes the coarser resolution pixels highly mixed. In a similar way, as the scale factor increases the optimal λ_p value decreases. This implies that, the contribution of the likelihood energy from the panchromatic image decreases. Therefore, to obtain better classification accuracy the likelihood energy from the multispectral band should be high.

As can be seen from Figure 4.1, when class separability between spectral classes increases the optimal λ value also increases. This result pointed out that, for higher class separability between spectral classes, the global energy needs to be modeled with a higher λ value. Whereas, in the case of lower separability of classes the global energy has a tendency to depend on the likelihood more than the prior energy. In other words, for lower class separability of classes, the class labelling depends more on the spectral information. This substantiates the need for accurate estimation of mean and covariance matrices for each spectral class, to obtain better classification results.

The classification accuracy kappa value decreases with increasing scale factor as shown in figure 4.5. However, an exception from this observation is an increase in classification accuracy with increasing scale factor value for class separability $JM^{(y)} = 0.5$ and $JM^{(z)} = 0.02$ at scale factor 5. The reason may be due to the less number of realizations of the images. With the same scale factor and class separability information the classification accuracy kappa value is depends on the optimal parameters of λ and λ_p as can be seen in figure 4.5.

The models relate λ and λ_p parameter estimation to scale factor and class separability information of the multispectral and panchromatic image. To estimate the parameters with the models only scale factor and class separability information is needed. The estimation is done by considering minimal class separability and an average class separability between spectral classes. The results achieved with these models for the estimation of the optimal parameters agree with the experimental data as shown in Figure 4.6, 4.7, 4.8.

The models are developed based on the assumption that the covariance matrices of the spectral classes are the same. This assumption simplifies the modeling significantly and considering unequal covariance matrices will make the model more powerful and realistic. The synthetic image which is used in this study was constructed with the assumption that the fine resolution pixels are conditionally independent. This assumption simplifies the construction of the synthetic images. However in real images, the spectral values

of pixel are spatially correlated. The applicability of the models to real image is tested and promising result was found however the models should consider the spatial structure of the pixels to be realistic. To apply the proposed models for real images first the scale factor value is decided, which is used as an input for SRM as well as for the model. Then the land cover classes are defined. Using the training set the class parameters mean and covariance of each class from the multispectral and mean and variance from the panchromatic band is estimated. These values are used for calculating the class separability (Bhattacharyya distance) between spectral classes in the multispectral and panchromatic image. Finally the optimal λ and λ_p parameter are estimated with the model based on class separability and scale factor information.

5.2. Conclusion

Automatic MRF parameter estimation is a challenging task, which is addressed in this research. The models are developed based on local energy balance analysis for the automatic estimation of the optimal parameters λ and λ_p in MRF based SRM. The Models relates the parameter estimation to scale factor and class separability information. The results of the experiments carried out using the synthetic images illustrate that the optimal λ and λ_p parameters depend on the scale factor and class separability. When the scale factor increases the optimal λ and λ_p parameters decreases. On the contrary, when the class separability increases the optimal λ and λ_p parameters also increase. Furthermore, for each combination of scale factor and class separability values in the multispectral and panchromatic images, an optimal λ and λ_p values exist. There exists also a range of λ and λ_p values for which the classification accuracy is close to the optimal value. The parameter values estimated with the model agreed with the results obtained from the experiment. The developed model allows users to estimate optimal λ and λ_p values by using the scale factor and class separability information of the panchromatic and multispectral images. Because of the incorporation of the panchromatic band, the result obtained for the optimal smoothness parameter λ has a slight difference with the result obtained in Tolpekin and Stein (2009)

5.3. Recommendations

1. It is recommended to improve the models for real images where the spectral values of pixels are Spatially correlated.
2. It is also recommended to improve the models by considering unequal class covariance matrices of the Spectral classes.

LIST OF REFERENCES

- Hailu Kassaye, R. (2006). *Suitability of Markov random field based method for super resolution land cover mapping*. ITC, Enschede.
- Jia, X., and Richards, J. A. (2008). Managing the Spectral-Spatial Mix in Context Classification Using Markov Random Fields. *Geoscience and Remote Sensing Letters, IEEE*, 5(2), 311-314.
- Kasetkasem, T., Arora, M. K., and Varshney, P. K. (2005). Super-resolution land cover mapping using a Markov random field based approach. *Remote Sensing of Environment*, 96(3-4), 302-314.
- Li, S. Z. (2001). *Markov random fields modeling in image analysis*. Tokyo etc.: Springer.
- Richards, J. A. (1993). *Remote sensing digital image analysis : an introduction* (Second revised and enlarged edition ed.). Berlin etc.: Springer-Verlag.
- Serpico, S. B., and Moser, G. (2006). Weight Parameter Optimization by the Ho–Kashyap Algorithm in MRF Models for Supervised Image Classification. *Geoscience and Remote Sensing, IEEE Transactions on*, 44(12), 3695-3705.
- Solberg, A. H. S., Taxt, T., and Jain, A. K. (1996). A Markov random field model for classification of multisource satellite imagery. *Geoscience and Remote Sensing, IEEE Transactions on*, 34(1), 100-113.
- Swain, P. H. e., and Davis, S. M. e. (1978). *Remote sensing : the quantitative approach*. New York etc.: McGraw-Hill.
- Tatem, A. J., Lewis, H. G., Atkinson, P. M., and Nixon, M. S. (2002). Super-resolution land cover pattern prediction using a Hopfield neural network. *Remote Sensing of Environment*, 79(1), 1-14.
- Tolpekin, V. A., Ardila Lopez, J. P., and Bijker, W. (2010). Super - resolution mapping for extraction of urban tree crown objects from VHR satellite images. In: *GEOBLA 2010 : geographic object - based image analysis, 29 June-2 July 2010, Ghent, Belgium : proceedings / editor E.A. Addink, F.M.B. Van Coillie. - [s.l.] : International Society for Photogrammetry and Remote Sensing (ISPRS), 2010. 7 p.*
- Tolpekin, V. A., and Stein, A. (2009). Quantification of the effects of land - cover - class spectral separability on the accuracy of Markov - random - field - based superresolution mapping. *IEEE Transactions on Geoscience and Remote Sensing*, 47(9).
- Tso, B., and Mather, P. M. (2001). *Classification methods for remotely sensed data*. London etc.: Taylor & Francis.

APPENDIX

A. Summary of results for optimal λ and λ pan values.

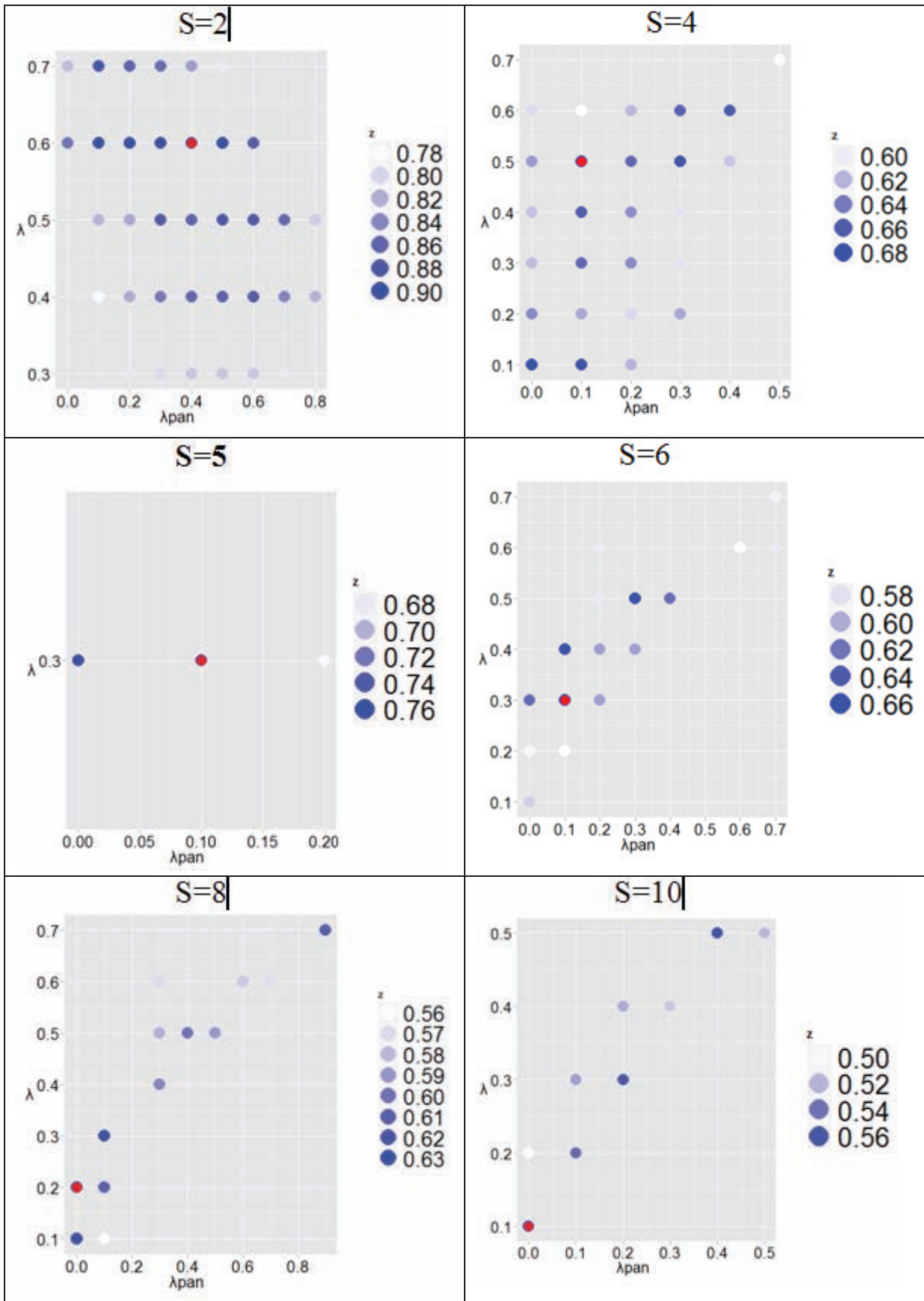


Figure A.1: Combination of λ and λ_{pan} values that gives kappa value above $0.85K_{max}$ at $JMY=0.5$ and $JMZ=0.02$

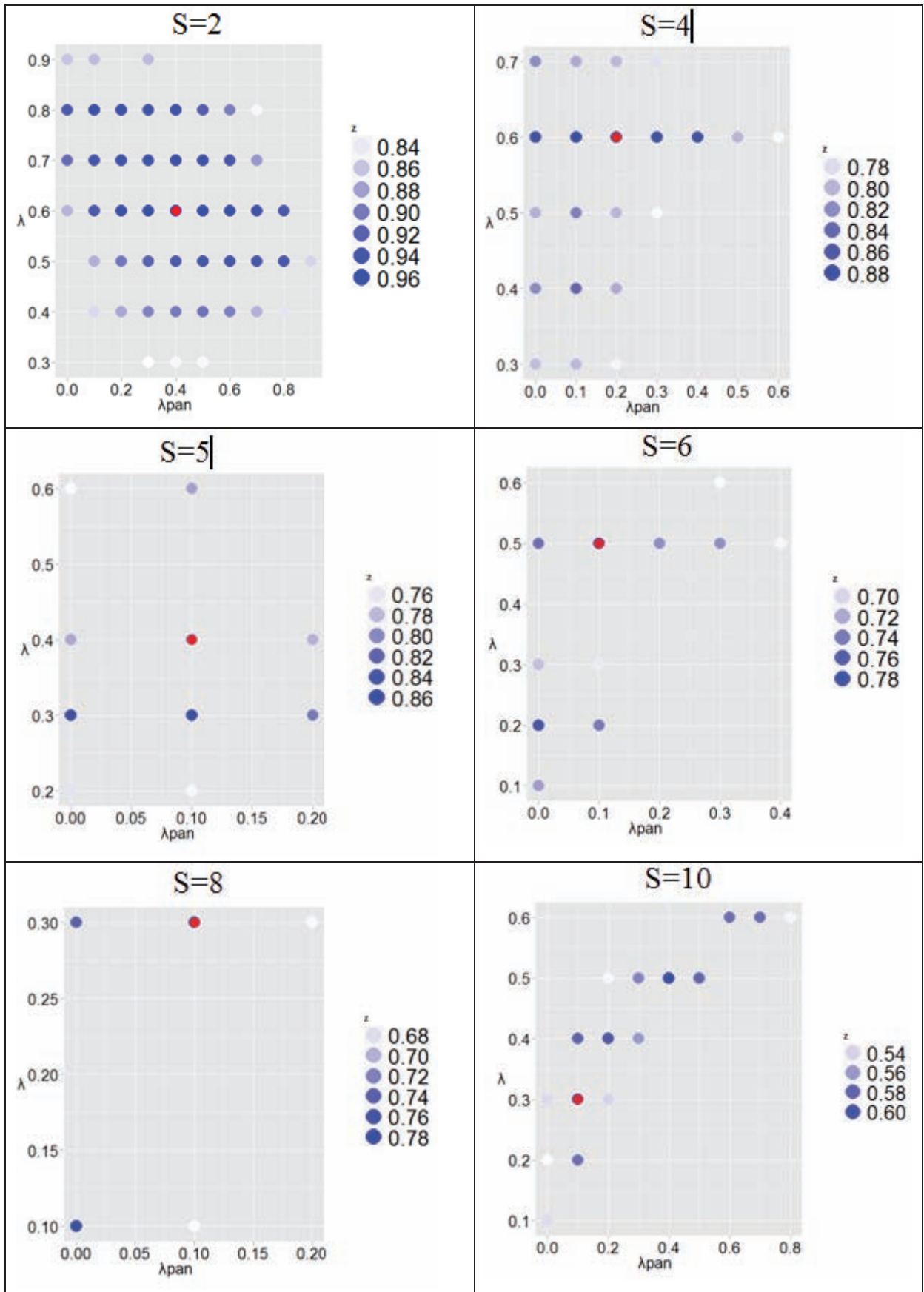


Figure A.2: Combination of λ and λ_{pan} values that gives κ value above $0.85K_{max}$ at $JMY=1.0$ and $JMZ=0.02$

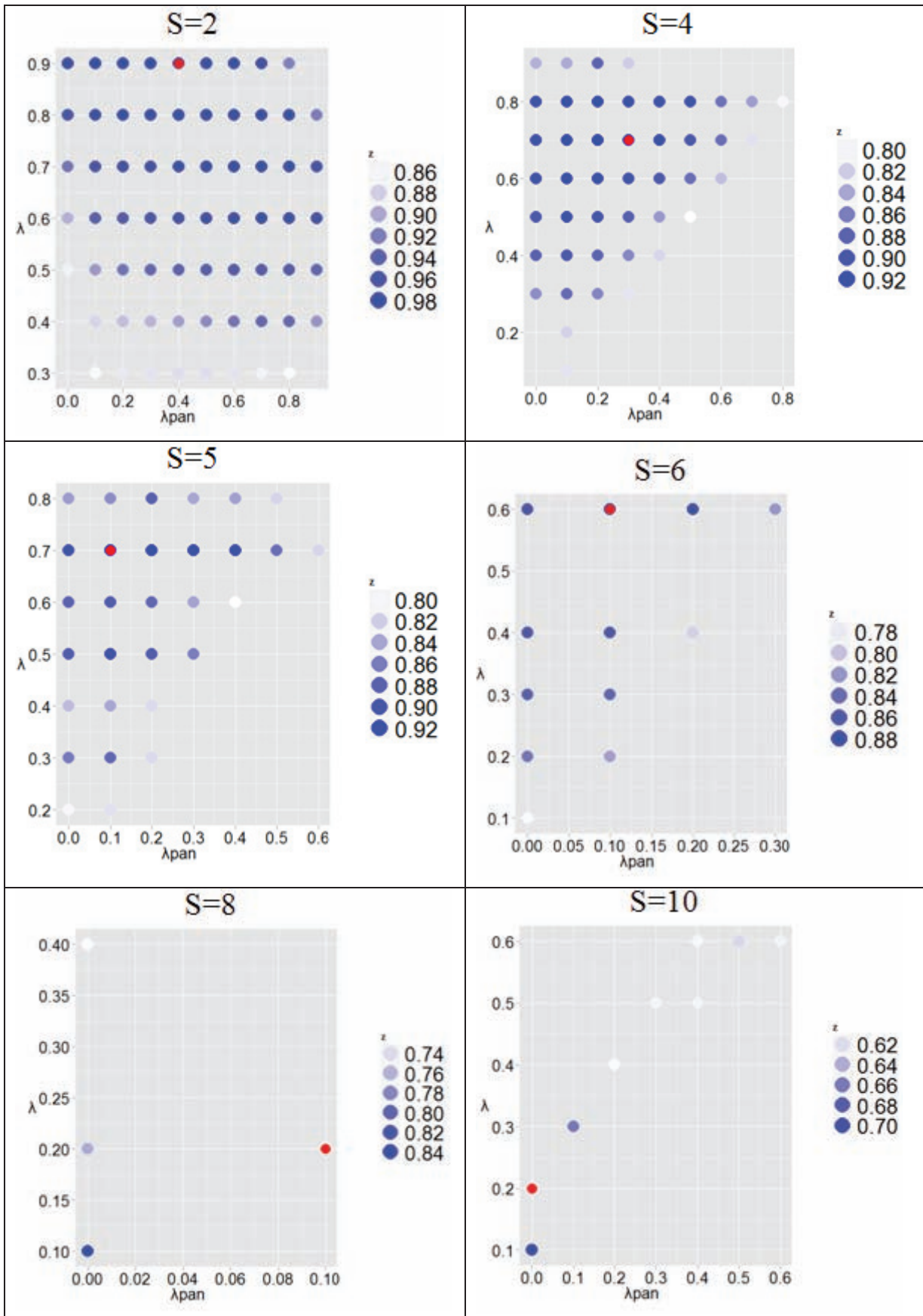


Figure A.3: Combination of λ and λ_{pan} values that gives kappa value above $0.85K_{max}$ at $JMY=1.9$ and $JMZ=0.02$

B. Summary of the result for optimal λ and λ pan estimation with average class separability.

scale	<i>JMY</i>	<i>JMZ</i>	$\hat{\lambda}^*$ pan	$\hat{\lambda}^*$
2	0.50	0.02	0.14	0.98
2	1.00	0.02	0.15	0.98
2	1.90	0.02	0.45	0.98
2	2.00	0.50	0.50	0.99
2	2.00	1.00	0.50	0.99
2	2.00	1.90	0.60	0.99
4	0.50	0.02	0.16	0.93
4	1.00	0.02	0.15	0.94
4	1.90	0.02	0.42	0.95
4	2.00	0.50	0.50	0.99
4	2.00	1.00	0.50	0.99
4	2.00	1.90	0.50	0.99
5	0.50	0.02	0.12	0.90
5	1.00	0.02	0.15	0.91
5	1.90	0.02	0.17	0.92
5	2.00	0.50	0.50	0.99
5	2.00	1.00	0.50	0.99
5	2.00	1.90	0.50	0.99
6	0.50	0.02	0.11	0.88
6	1.00	0.02	0.14	0.88
6	1.90	0.02	0.15	0.90
6	2.00	0.50	0.40	0.98
6	2.00	1.00	0.40	0.99
6	2.00	1.90	0.60	0.99
8	0.50	0.02	0.09	0.82
8	1.00	0.02	0.12	0.83
8	1.90	0.02	0.13	0.85
8	2.00	0.50	0.40	0.98
8	2.00	1.00	0.40	0.99
8	2.00	1.90	0.60	0.99
10	0.50	0.02	0.07	0.77
10	1.00	0.02	0.10	0.78
10	1.90	0.02	0.11	0.80
10	2.00	0.50	0.40	0.97
10	2.00	1.00	0.40	0.99
10	2.00	1.90	0.60	0.99