

Multi sensor data and temporal image fusion cross validation technique for Agri yield monitoring system

Swati V. Shinde (✉ swaatii.shinde@hotmail.com)

Pimpri Chinchwad College of Engineering

Rajveer Shastri

Vidya Pratishthan's College of Engineering Baramati: Vidya Pratishthan Vidya Pratishthans Kamalnayan Bajaj Institute of Engineering and Technology Baramati

Atul Kumar Dwivedi

Bundelkhand Institute of Engineering and Technology

Anandakumar Haldorai

Sri Eshwar College of Engineering

Varsha Sahni

CT Institute of Engineering Management and Technology

Balaji Adusumalli

Guntur Engineering College

Research Article

Keywords: Multi-temporal and multimodal data fusion, cross validation technique, yield monitoring system, multimodal datasets

Posted Date: September 30th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-943821/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

In recent years, the diverse application in various disciplines and the versatility has gained a huge interest for the researchers to research on the multi-sensor data fusion technology. The remote sensing process involves the measurement and recording of the data from a scene. Thus, the remote sensing systems are known to be a powerful tool as they help in the earth's atmosphere and surface monitor at different scales. The remote sensing of the data faces a serious challenge as the data captured by the multiple sensors are heterogeneous. This affects the efficient processing and the effectiveness of the data that is being sensed. Thus, the increase in the diversity in data increases the ancillary datasets. These multimodal datasets are used jointly to improve the processing performance as per the application requirement. Initially, the fusion of the temporal data with the backscattered/temporal data is possible from the data retrieved from remote sensing. Many researchers made several types of research on fusing the multi-temporal and multimodal data and gave different ideas for a different type of researchers. This paper presents the cross-validation technique for monitoring the yield. This monitoring system is developed by fusing the multi-sensor data and the temporal images. This fusion is performed, and the performance of the yield monitoring system is analyzed from the results obtained. By using the cross-validation technique, the efficiency of the system is found to be improved.

1 Introduction

Remote sensing, the methodology that is being used in a wide range of applications like civilian applications, military, and surveillance C. Elachi (1987). The remote sensing process involves the measurement and recording of the data from a scene. Thus, the remote sensing systems are known to be a powerful tool as they help in the earth's atmosphere and surface monitor at different scales. The scales vary from global to local scale. These monitoring systems can provide vital mapping and coverage and the land coverage classifications like a forest, soil, water bodies, vegetation, etc. The grade of the accuracy greatly depends on the following: 1. Knowledge of the researchers, 2. Area cover types and 3. Image quality. The relation between the forest, soil, and the forest is obtained when there is a correlation plotted among the superficial deposits, drainage, and features of the topography. The above details become vital for the management of the land in use and the classification of the land topography. The air/space borne sensors are being used in the process of remote sensing. These sensors can acquire the data at varying spectral bands on the oppressed frequency basis at numerous resolutions. The data related fusion aspects is shown in Fig. 1.

There are numerous data available at the wide spectrum from the sensors for the same scene or the site observed. The single-unit sensor's data becomes inconsistent, incomplete, and even imprecise for many applications ((P.K. Varshney (1997); (D.L. Hall et al.1997); (C. Pohl et al.1998)) The uncertainty and the errors due to single-unit sensor information may be decreased with the help of multi-sensor data as they provide compatible data which are used in the fusion for giving the observed site a better understanding (A. Farina et al. 1996; A. Farina et al. 1996;V. Clement et al.1993). The remotely sensed images are provided either by radar or by heterogeneous sensors. Thus, the multi-temporal images provided for the

same site are needed to determine the changes that had occurred in the considered site. The site segment classification needs the multi-dimensional images of the radar and non-radar sensors as the segment separation increases because of the multi-polarization and multi-spectral data (A.Farina et al. 1984; F.W. Leberl et al. 1990). The vital information in the interpretation of the site scene is the contextual data. When it comes to labeling, the pixel in isolation provides incomplete information of the characteristics that are wanted. Contextual information is mainly dependent on the three domains: 1. Time, 2. frequency, and 3. Space. The different electromagnetic bands of the spectrum are known as the spectral dimensions. These dimensions are either obtained by a single-unit sensor or by the multiple frequencies operating numerous sensors. The single band images offer less resolution compared to the separation of various ground cover classes. The complementary data of the same observed site are available for the image fusion as follows: 1. Multi-sensor image fusion, 2. Multi-temporal image fusion, 3. Multi-frequency image fusion, 4. Multi-polarization image fusion, and multi-resolution image fusion. The remaining sections of the paper are organized as follows. Section 2 contains the description of the detailed study made related to our work. Section 3 contains a detailed description of the components used in the yield monitoring system. Section 4 contains the methodology used for the data fusion of the temporal images and its data processing, classification algorithms. This section also provides information about difficulties involved in the classification of the data. Section 5 contains the experimental results and their discussion.

2 Related Work

The remotely sensed data from the satellite platform is of different bandwidth and varying spatio-temporal resolutions. Thus, these data from the satellite are always available for the requirement of having both temporal and spatial resolutions high with a single observation (B.Chen et al.2015). There was a shallow recess between the data fusion and data merging. This helped in the improvement of the resolution over the spatial or temporal data. The required information for the data retrieval in the remote sensing applications are 1. LSR (Land Surface Reflectance), and 2. (TAR) Top of Atmosphere Reflectance. The above algorithms are used in the retrieval of data in many fields which include Land cover types (Y.Ke et al.2016) thermal IR radiance monitoring (Q.Wang et al. 2017), evapotranspiration mapping (C.Kuenzer et al. 2013), crop yields (T.Kleynhans et al.2017), and atmospheric carbon balance (K. J. Lees et al .2018). Under complex environmental conditions in the urban areas, the public health assessment and the air quality management needs accurate information from both LSR and TAR. These data are also needed for aerosol optical depth retrieval. The derivation of the data from the surface Reflectance at the top of the atmosphere and the surface land levels with an improvement in the temporal resolution and the spatial coverage (C. Pohl et al.1998). The improvement in the prediction of reflectance data for higher-end applications with a fine-tuned Spatio-temporal (R. Lunetta et al.1998) resolution helps not only in the advancement but also to avoid physical disadvantages (Z.Wang et al. 2017).

The remote sensing of the modern era is integrated with data fusion and data margin, as there is a salient gap between the sensor's random factors which includes cloud contamination platforms A. C. Povey et al.2015; N. Bin et al.2018). To overcome the salient gaps between various data, data fusion (K. Bai et

al.2020), data margin (L. Wald et al.1999) and data reconstruction (K. Bai et al.2016; H. Nguyen et al.2017) have been fused in the past. There is a need for a huge amount of data that are spatiotemporally adjacent or the historic data in the pixel removal and lost data reconstruction (A. Chatterjee et al. 2010). A similar design of the two distinct satellite platforms can be used in the data merging process which helps partially in the retrieval of the lost data. The spatial, temporal, and spectral images can be enhanced using the well-known technique called data fusion (J. Dong et al.2009). Data fusion tends to extract data from asynchronous time series of the satellite data whereas the data merging algorithms uses the synchronous time series data from the satellite with varying temporal, spectral, and spatial resolution. The data fusion technique enhances the spectral, spatial, and temporal properties and converts them into a synthetic image (J. L. Van Genderen et al.1994).

The widely used and attention gained data fusion traditional methods are 1. The STARFM (spatial and temporal adaptive reflectance fusion model), and BME (Bayesian maximum entropy). But in the recent years of improvement in the data fusion technique, the ML (machine learning) algorithms have been used for their promising results in particular the problems related to the environment X. Zhu et al. 2018; P. Dai et al.2018; H. Song et al.2018). The widely used algorithms in the data fusion technique are 1. Extreme ML (C. He et al.2020), 2. Dictionary learning (J. Li et al.2020), 3. Random forest 4. Regression tree (S. Boyte et al.2017), and 5. Artificial neural network. The ML-based algorithms for the fusion of the Spatio-temporal data are limited with the creation of effective data training (Y. Liu, et al.2018), as there are countless parameters available (G. Li et al.2020). This leads to less accuracy in the prediction due to the resolution differences across the various image sources (S. M. Howard. 2017)

The input data uncertainty can be well handled by using Bayesian maximum entropy to showcase its efficacy in the fusion of data. There is no elevated requirement needed for the input image numbers and it also had good potential in handling dataset of temporal resolution (J. He et al.2019). The extension of the STARFM which is widely used in the invariant temporally in nature is 1. The ESTARFM (enhanced spatial and temporal adaptive reflectance fusion model) (T. Hilker et al. 2009), 2. The SPSTFM (sparse representation-based spatiotemporal reflectance fusion model) (X. Zhu et al 2010) 3. The ESTARFM (enhanced spatial and temporal adaptive reflectance fusion model) (T. Hilker et al 2009), 4. The ESTDFM (enhanced spatial and temporal data fusion model) (C. Cammalleri et al.2014), and 5. The SPSTFM (sparse representation-based spatiotemporal reflectance fusion model) (F. Zhang et al.2014).

The vital factors that affect the development of the data fusion or the data merging algorithms irrespective of the land surface or the top of the atmosphere environment are 1. characteristics of the illumination, 2. satellite platform's spatial and spectral properties, 3. contamination in the cloud, and angle of view (C. Song et,2003). The different parameters of the land surfaces are robustly measured using the different parameters with the help of the newly designed approach which provides multi-sensor data along with the data fusion. These approaches are capable of providing accurate, complementary, and robust output from the remotely sensed data. There are three sensor fusion techniques and they are remote with proximal sensor fusion, proximal with proximal sensor fusion, and remote with remote

sensor fusion. The data fusion integrates the data from the multi-sensors and the accuracy in the output is more precise than the data from the single sensor output.

The right technique chosen for the measurement greatly influences the measurement accuracy. Choosing the sensor set to be integrated is greatly dependent on the following: 1. Practical information, 2. Fusibility of the sensor, 3. Objective parameters, and 4. Requirement and situation of the application. There are two types of sensors in which one type is easily available for blend whereas the other type needs a higher-end calibration, regular monitoring of the data interpretation and processing.

3 Proposed System Model

The proposed system model is mainly depending on the volumetric flow measurement. The various sensors are engaged in the enhancement of the yield monitor. The system model includes the GPS, junction box, optical sensor, and field computer. The flowchart of the multi-sensor data and temporal image fusion is shown in Fig. 2. At the top of the clean grain elevator, an optical sensor is placed beside it to calculate the grain yield. The elevator housing includes the fixed hinged mounting bracket that encloses the transmitter and receiver in association with the lens and lens holders. The operational function of each sensor is exhibited via the light-emitting diode. From one to the other end of the elevator paddles, the infrared light beam is transmitted. If there is any interruption in the light beam is detected using the receiver. When the paddle allows the transmission of the sensor across it, the beam is broken. The beam breaks for a longer period depending on the amount of grain available on the paddle.

The Global Positioning System (GPS) is a sensor-based yield monitoring system. The satellite signals are utilized in the proposed monitoring system. The main purpose of the sensor-based GPS is to identify the desired location and in displaying the combined speed via satellite images. The main benefit of GPS mapping technology is monitored using the sub-meter accuracy. The association of the ground-related segments with the space is well-known as Global Positioning System. Above the ground level, at the high central point, GPS is usually mounted.

The position of the field computer is usually opposite to the driver seat. The output signals of the sensor are stored and displayed by the field computer where these signals are considered as data. The signal obtained from the sensors are later used in the field computer includes the GPS interface, user graphical user interface, external data storage devices, and controls the communication among various devices. The field computer can perform various functions such as sensor integration and calibration, moisture content, online speed, the area covered, and yield per hectare. The yield sensor, GPS receiver, and moisture sensor header data are collected in the Junction box. The Junction Box also records the header height, cut width, and GPS receiver. The correction factor, crop type, field name, and calibration number are to be filed in the Junction box. The Junction box is placed inside the cab mainly for safety and protection purposes rather than using as weather proof. The path for the sensor cables is attained through the bulkheads depend on the mounting position of the Junction box.

The header of the combine encloses the on/off switch of the spring. The association of the large number of the harvester in the fieldwork is differentiated using the yield monitor switch. When the header is in an upward position then the data logging is retained. The Junction box starts to count the area when the switch is in ON condition. The switch is ON condition when the combine header is in working position.

The temporal phenomena are used in the analysis of the remote sensing data by the time variable. The satellite constellation namely, Sentinel 1 and Sentinel 2 includes the features such as fine resolution images and short revisit time. These characteristic features make the satellite remote sensing images obtain the semantic content from a scene. In space and time, the dimensionality changes from the 3D to 4D by the time variable. The time variable in association with the image pair, short time-series, or long time-series SAR active images are exploiting the performance of the time variable. In the structural multi-temporal sensor-fusion, time-series multi-sensor images are fused. The perspective point-of-view of various techniques in the analysis of data is obtained via the fusional temporal information along with the image spatial, image backscattering, and image spectral information.

The multi-temporal data includes the analysis of image time series. The fine spatial resolution images are acquired through the association of spectral, spatial, and temporal informational data on dense time-series. Sentinel 1, Sentinel 2, and Landsat-8 are the few dense time-series that produce fine resolution images which can be further improved. These data are used in the analysis of the field vegetation in the agricultural farms to improve their contemporaneous applications. An individual multi-sensor time-series obtained from the various satellite are fused is also capable of providing precise farming. Depending on the consequence and its extension, the analysis of the temporal information varies. Initially, the consequence is classified concerning the multi-temporal data and the solution of the respective consequence is attained later. The labeled training data related to the multi-temporal classification challenges are encountered.

The multi-temporal images for at least a pair are obtained from the geographical area at various times. Based on the analysis of data objective, temporal data are classified into various types. The recent images of the land cover map with the time-series, multi-temporal land cover maps set are attained from each item of the time-series, and every pixel in images is represented as temporal information of the seasonal land cover map. Depending on the illustration, the multi-temporal is classified into three different ways. In an application-oriented system, the identification of the problem is very difficult to implement.

The representation of the multi-temporal data stacked vector is generated as an input to the classifier in the supervised direct multi-date classification. The image stacking vector can be acquired by several times with the characterized pixels. The newly available images produce the map regarding the land cover concerning the training classification. Image acquisition dates measured based on the land cover remain unaltered, and the data distribution model concerning the various methodologies. The statistical Bayesian method is utilized for evaluation in later times. Each attribute of the time-series land cover is categorized using the multi-date direct classification method. Therefore, currently, the present acquisition

time for the corresponding land cover map is obtained. Thus, the transition of the explicit land cover can be determined and the removal of the assumption in between the selected data will not affect the progress. The probability of class combined with the adequate training data provides the information for the framework of the proposed model. The proposed model is applicable in contemporaneous applications. To overcome the consequence in a multiday direct classification, various methodologies are adopted.

The cascade categories of image pairs are involved in classifying the multi-temporal data. The block scheme of the temporal classification is shown in Fig. 3. At different periods, the temporal correlation is utilized to link the probability of classes in a single image based on the Bayesian perspective. The proposed model consists of the distribution class for all individual data to determine the temporal correlation in between the images. The proposed multi-temporal and multi-sensor data includes the neural network classifier in association with the Bayesian decision framework. The proposed structure is free from the distribution estimation as it is acquired from the multispectral and SAR multi-temporal images.

The fusion model is usually preferable in the classification scheme. Kernel methods and multiple classifier systems and several neural models namely, radial basis function networks, and multilayer perception neural networks are used in the fusion model. The challenges in the classification of multi-temporal data using the deep learning architectures namely convolutional neural networks are minimum. Thus, the proposed model framework apprehends the Spatio-temporal arrangement. The precise rate over the land-cover-transition maps is high.

The multi-temporal image development is achieved through the deep learning framework with the affordable computation is the significant drawbacks. The deep learning framework in association with the 4D data structure is more complex and hence, a large amount of training data is required for effective performance. The remote sensing data is categorized depending on the availability of labeled samples with the time information sources. These data are trained to learn the supervised algorithms.

The temporal-based modeling includes multi-temporal classes, the time-series-based linkage between the various classes, and distinct spatial with the high temporal value. The semi-supervised classification methods are achieved via the combination of labeled training data with the recorded images to develop the proposed model. The spatial-temporal properties include the estimation of time-series. The remote sensing scheme makes use of the expectation-maximization algorithm regarding the land-cover map. This scheme is applicable in the classification of the cascade, compound, and bi-temporal images. Thus, the proposed system model consists of the linked multispectral, SAR multi-temporal, and multi-sensor images. The proposed system utilizes the active learning structural compound classification in improving the training data with a low-cost collection.

The high multi-temporal label classes in the images are used to acquire ad-hoc training data samples. The proposed transfer learning scheme performs the propagation function with the data of the given image to the training sets of other images with the constrained time-series. The class labels of the

images are transmitted within the constrained time-series to the remaining pixels which remain unaltered. The multi-temporal classifier is provided with the increased supervision due to the changes that occurred on the unsupervised detection. In the classification of multi-temporal data, a partially supervised 4D data framework along with the deep learning architecture is provided. Finally, the decoupling of the network is achieved in relationship with the spatial-temporal pattern without affecting the extraction capability.

4 Result And Discussion

The growing and the increasing availability of various sensor captured data connected with the help of the computational tools and the methodological approaches make them attractive in the fusion of the diversified datasets that are complementary. This helps in increasing the data processing system's efficiency and capacity of the remotely sensed data concerning the problems that are available. This detailed interpretation, it is shown clear that the data fusion technique is achieving more importance and popularity. There are few applications of the data fusion technique, which will not be given importance in the future. The reports and documents generated by the humans will act as the source for the future technology. The powerful frameworks will be developed, and they will have the capacity to process wide variety of data.

The Fig. 4 shows the result of the simulation showing the root mean square error distribution (RMSE). This result is obtained when there is a prediction of a trait rather than replacing it from the random and blocked cross-validation with different values for k . The experiment provided the result which is 5–95% true value RMSE range of the real model.

The Fig. 5 provides the forecast estimate by plotting the cross-validation that is being repeatedly performed varying the block size. This result is from the performance threshold which is drawn. For a performance prediction threshold given, the projection requirement is varied. Thus, a range of values is held by the forecast horizon. The forecast horizon is marked in such a position where after that the projection is too useless and unreliable.

5 Conclusion

This paper presented a multi-sensor data and temporal image fusion, cross-validation and classification technique and algorithms. The data fusion is an art methodology state. The data fusion has a wide application range, and it is known to be the multi-disciplinary field for research. We have developed a data centric arrangement for the methodology of the data fusion. This system is designed to explore the theoretical framework and their associated aspects that are challenging. This paper also involves the introduction of algorithms for the existing categories. This paper also includes the exploration of the fusion technologies, it is vivid that the system of data fusion is becoming the common place application's technology. There is an increased demand for the large-scale data fusion technology which involves the web and the sensor networks. Thus, the highly scalable algorithms for the data fusion are demanding an increased interest on distributed architectures. As the result, there is a trend developed which will increase

the adoption of the scalable data fusion algorithms in the future. This trend also motivated the researchers to perform more research on the topics related to the performance of the data fusion systems in real time applications for security and reliability purposes.

Declarations

Conflict of interest : The authors declare that they have no conflict of interest.

Funding : Not Applicable

Availability of data and material: Not Applicable

Code availability: Not Applicable

References

1. Elachi C (1987) Introduction to the physics and techniques of remote sensing, in: J.A. Kong (Series Ed.), Wiley Series in Remote Sensing
2. Varshney PK ((1997), Multisensor data fusion, Electron. Commun. Eng. J. 245–263
3. Hall DL, Llinas J (1997), An introduction to multisensor data fusion, Proc. IEEE 85 (1) 6–23
4. Pohl C, van Genderen JL (1998) Multisensor image fusion in remote sensing: concepts, methods, and applications. Int J Remote Sensing 19(5):823–854
5. A.Farina M, Costantini F, Zirilli (1996) Fusion of radar images: techniques and applications, Invited Lecture, Colloquium Entretiens Science et Defense, Topic on Le futur du radar: une synthesis de techniques 24–25, January 1996, 285–296
6. A.Farina M, Costantini F, Zirilli (1996) Fusion of radar images: techniques and applications, Mil. Techn Miltech 5:34–40
7. Clement V, Giraudon G, Houzelle S, Sandakly F (1993) Interpretation of remotely sensed images in a context of multisensor fusion using a multispecialist architecture, IEEE Trans. Geosci. Remote Sensing 31 (4)
8. A.Farina FA, Studer (1984) Radar Data Processing, vol 1. Research Studies Press, Wiley, New York
9. F.W. Leberl(1990), Radargrammetric image processing, Artech House
10. Chen B, Huang B, Xu B (2015) Comparison of spatiotemporal fusion models: A review. Remote Sens 7(2):1798–1835
11. Y.Ke J, Im S, Park, Gong H (2016) Downscaling of MODIS one kilometer evapotranspiration using Landsat-8 data and machine learning approaches. Remote Sens 8(3):1–26
12. Q.Wang et al. 2017, “Fusion of Landsat 8 OLI and Sentinel-2 MSI data,” IEEE Trans. Geosci. Remote Sens., vol. 55, no. 7, pp. 3885–3899, Jul.

13. C.Kuenzer and Knauer K(2013), "Remote sensing of rice crop areas," *Int. J. Remote Sens.*, vol. 34, no. 6, pp. 2101–2139, 2013
14. T.Kleynhans M, Montanaro A, Gerace, Kanan C (2017) Predicting top-of-atmosphere thermal radiance using MERRA-2 atmospheric data with deep learning. *Remote Sens* 9(11):1–16
15. Lees KJ, Quaife T, Artz RRE, Khomik M, Clark JM (2018) Potential for using remote sensing to estimate carbon fluxes across northern peatlands – A review. *Sci Total Environ* 615:857–874
16. Pohl C, Van Genderen JL (1998) Review article Multisensor image fusion in remote sensing: Concepts, methods and applications. *Int J Remote Sens* 19(5):823–854
17. Lunetta R, Lyon J (1998), "North American landscape characterization dataset development and data fusion issues," *Eng Remote Sens*, 64, 8, 821–829
18. Z.Wang S, Yang Q, Zeng, Wang Y (2017) Retrieval of aerosol optical depth for Chongqing using the HJ-1 satellite data. *J Meteorol Res* 31(3):586–596
19. Povey AC, Grainger RG (2015), "Known and unknown unknowns: Uncertainty estimation in satellite remote sensing," *Atmos. Meas. Tech.*, vol. 8, no. 11, pp. 4699–4718, 2015
20. Bin Chang N, Bai K, Imen S, Chen CF, Gao W (2018) Multisensor satellite image fusion and networking for all-weather environmental monitoring. *IEEE Syst J* 12(2):1341–1357, Jun
21. Bai K, Li K, Guo J, Yang Y, Bin Chang N (2020) Filling the gaps of in situ hourly PM_{2.5} concentration data with the aid of empirical orthogonal function analysis constrained by diurnal cycles. *Atmos Meas Tech* 13(3):1213–1226
22. Wald L (1999) Some terms of reference in data fusion. *IEEE Trans Geosci Remote Sens* 37(3):1190–1193
23. Bai K, Bin N, Chang, Chen CF (2016) Spectral information adaptation and synthesis scheme for merging cross-mission ocean color reflectance observations from MODIS and VIIRS. *IEEE Trans Geosci Remote Sens* 54(1):311–329
24. Nguyen H et al (2017) Multivariate spatial data fusion for very large remote sensing datasets. *Remote Sens* 9(2):1–19
25. Chatterjee A, Michalak AM, Kahn RA, Paradise SR, Braverman AJ, Miller CE (2010) A geostatistical data fusion technique for merging remote sensing and ground-based observations of aerosol optical thickness. *J Geophys Res Atmos* 115(20):1–12
26. Dong J, Zhuang D, Huang Y, Fu J (2009) Advances in multi-sensor data fusion: Algorithms and applications. *Sensors* 9(10):7771–7784
27. Van Genderen JL, Pohl C,(1994) "Image fusion: Issues, techniques and applications," in *Proc. EARSEL Workshop Intell. Image Fusion*, 1994, pp. 18–26
28. Zhu X, Cai F, Tian J, Williams TKA (2018) Spatiotemporal fusion of multisource remote sensing data: Literature survey, taxonomy, principles, applications, and future directions. *Remote Sens* 10(4):1–23
29. Dai P, Zhang H, Zhang L, Shen H,(2018) "A remote sensing spatiotemporal fusion model of Landsat and modis data via deep learning," *Int. Geosci. Remote Sens. Symp.*, vol. 2018-Jul., no. May 2019,

pp. 7030–7033

30. Song H, Liu Q, Wang G, Hang R, Huang B,(2018) "Spatiotemporal satellite image fusion using deep convolutional neural networks," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 11, no. 3, pp. 821–829, Mar. 2018
31. He C,(2020) "Spatio-temporal series remote sensing image prediction based on multi-dictionary Bayesian fusion," *Sci. China Inf. Sci.*, vol. 63, Art. no. 140301
32. Li J, Li Y, He L, Chen J, Plaza A,(2020) "Spatio-temporal fusion for remote sensing data: An overview and new benchmark," *Sci. China Inf. Sci.*, vol. 63, no. April, pp. 1–17
33. Boyte S, Wylie BK, Boyte SP, Wylie BK, Rigge MB, Dahal D,(2017) "Fusing MODIS with Landsat 8 data to downscale weekly normalized difference vegetation index estimates for central Great Basin rangelands, USA Fusing MODIS with Landsat 8 data to downscale weekly normalized difference vegetation index estimates for central Great Basin rangelands, USA," *GISci. Remote Sens.*, vol. 55, no. 3, pp. 376–399, 2017
34. Liu Y, Chen X, Wang Z, Wang ZJ, Ward RK, Wang X, (2018)"Deep learning for pixel-level image fusion: Recent advances and future prospects," *Inf. Fusion*, vol. 42, no. Nov. 2017, pp. 158–173
35. Li G, Liu Q, Zhao S, Qiao W, Ren X,(2020) "Automatic crack recognition for concrete bridges using a fully convolutional neural network and naive Bayes data fusion based on a visual," *Meas. Sci. Technol.*, vol. 31, no. 7, pp. 1–17, 2020
36. Howard SM,(2017) "Deep learning for sensor fusion," Thesis MS, Case Western Reserve Univ., Cleveland, OH, USA
37. He J, Christakos G, Jankowski P (2019) Comparative performance of the LUR, ANN, and BME techniques in the multiscale spatiotemporal mapping of PM_{2.5} concentrations in North China. *IEEE J Sel Top Appl Earth Obs Remote Sens* 12(6):1734–1747
38. Hilker T et al (2009) A new data fusion model for high spatial- and temporal-resolution mapping of forest disturbance based on Landsat and MODIS. *Remote Sens Environ* 113(8):1613–1627
39. Zhu X, Chen J, Gao F, Chen X, Masek JG (2010) An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. *Remote Sens Environ* 114(11):2610–2623
40. Hilker T et al.2009, "Generation of dense time series synthetic Landsat data through data blending with MODIS using a spatial and temporal adaptive reflectance fusion model," *Remote Sens Environ*, 113, 9, 1988–1999
41. Cammalleri C, Anderson MC, Gao F, Hain CR, Kustas WP,(2014) "Mapping daily evapotranspiration at field scales over rainfed and irrigated agricultural areas using remote sensing data fusion," *Agric. For. Meteorol.*, vol. 186, pp. 1–11, 2014
42. Zhang F, Zhu X, Liu D (2014) Blending MODIS and Landsat images for urban flood mapping. *Int J Remote Sens* 35(9):3237–3253
43. Song C, Woodcock CE (2003) Monitoring forest succession with multitemporal Landsat images: Factors of uncertainty. *IEEE Trans Geosci Remote Sens* 41(11):2557–2567

Figures

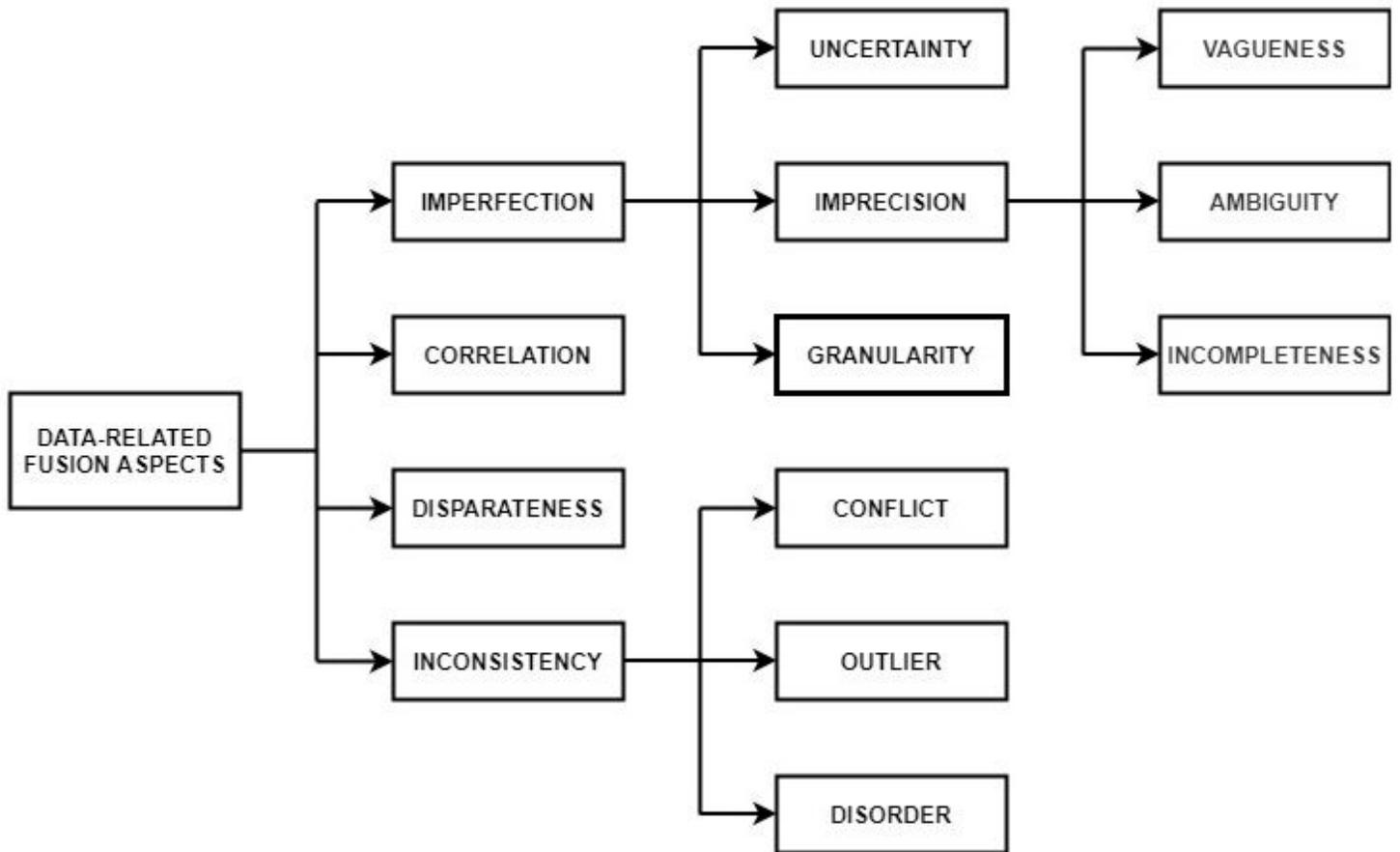


Figure 1

Characteristics of multi-temporal data related fusion.

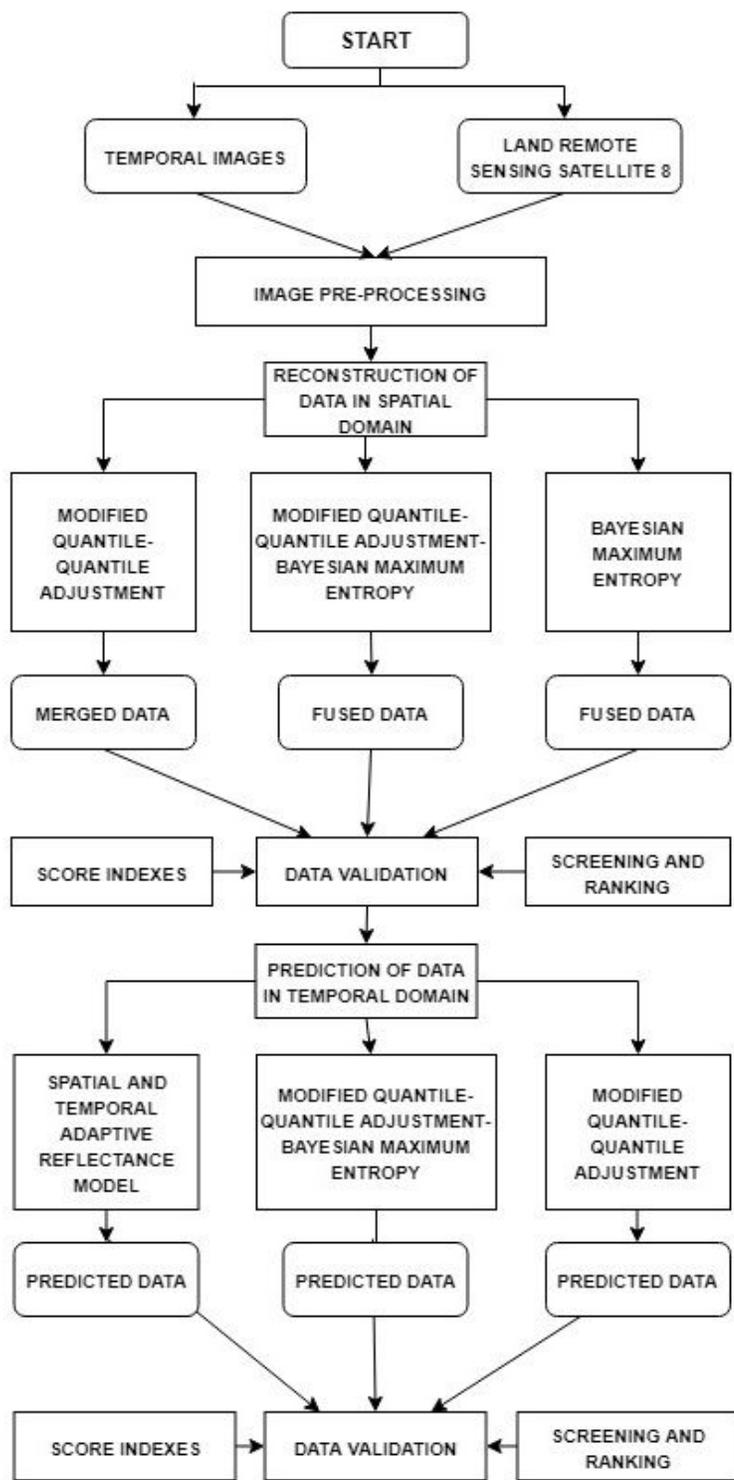


Figure 2

Flowchart for the multi-sensor data and temporal image fusion.

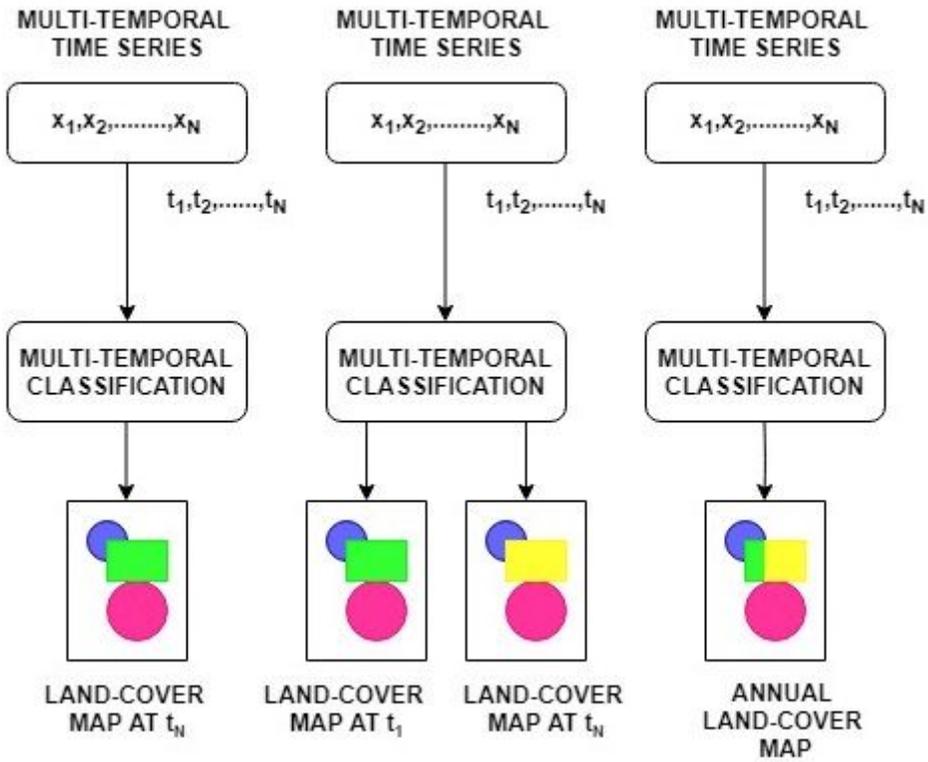


Figure 3

Block scheme of the temporal classification.

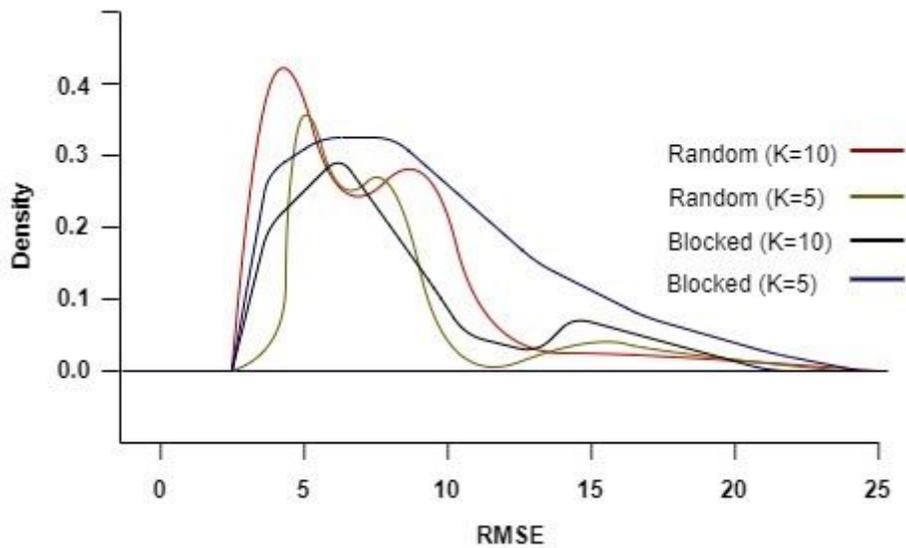


Figure 4

Root mean square error distribution (RMSE).

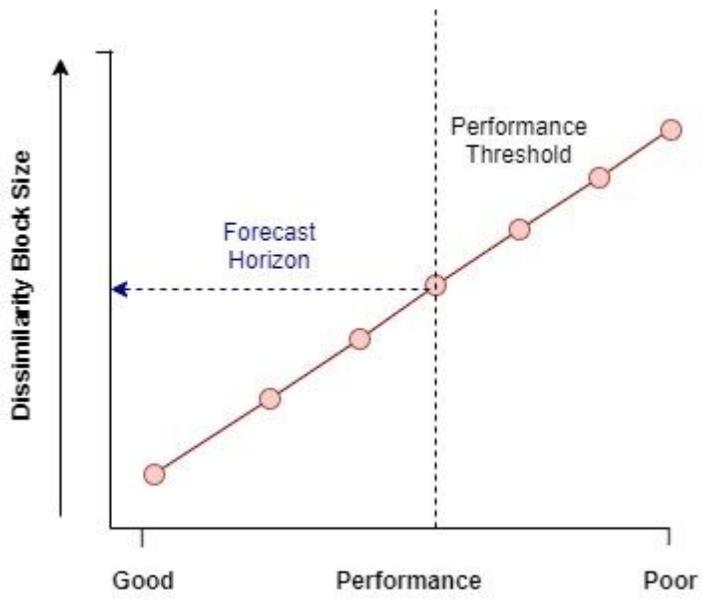


Figure 5

Performance evaluation for various block size.