

©2020 Society of Photo-Optical Instrumentation Engineers (SPIE). One print or electronic copy may be made for personal use only. Systematic reproduction and distribution, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited. Access to this work was provided by the University of Maryland, Baltimore County (UMBC) ScholarWorks@UMBC digital repository on the Maryland Shared Open Access (MD-SOAR) platform.

Please provide feedback

Please support the ScholarWorks@UMBC repository by emailing scholarworks-group@umbc.edu and telling us what having access to this work means to you and why it's important to you. Thank you.

PROCEEDINGS OF SPIE

SPIDigitalLibrary.org/conference-proceedings-of-spie

Satellite data fusion of multiple observed XCO₂ using compressive sensing

Nguyen, Phuong, Shivadekar, Samit, Chukkapalli, Sai Sree, Halem, Milton

Phuong Nguyen, Samit Shivadekar, Sai Sree Chukkapalli, Milton Halem, "Satellite data fusion of multiple observed XCO₂ using compressive sensing," Proc. SPIE 11423, Signal Processing, Sensor/Information Fusion, and Target Recognition XXIX, 114230Y (22 April 2020); doi: 10.1117/12.2558319

SPIE.

Event: SPIE Defense + Commercial Sensing, 2020, Online Only, California, United States

SATELLITE DATA FUSION OF MULTIPLE OBSERVED XCO₂ USING COMPRESSIVE SENSING

Phuong Nguyen^{*a}, Samit Shivadekar^a, Sai Sree Laya Chukkapallia^a, Milton Halem^a

^aDept. of Computer Science and Electrical Engineering, University Of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250

ABSTRACT

When it entered into the era of big data, Earth observing systems developed into a new stage, namely characterized by low cost, multi-national, multi-sensor and multi-modal with varying spatial and spectral resolutions confronting new challenges and opportunities. Climate data records from multiple data sources are used to infer seasonal and interannual variations which will advance and promote the development of data fusion methods. Compressed sensing is a new framework in which data acquisition and data processing are merged. It provides a new fantastic way to handle multiple observations of the same field view from complementary remote sensing instruments, allowing us to recover information at very low signal-to-noise ratio.

We will particularly point out that a Compressive Sensing based framework is flexible enough for combining the two measurement systems by fusing the data from the two satellites, NASA Orbiting Carbon Observatory -2 (OCO-2) and the JAXA Greenhouse gases from Orbiting Satellites (GOSAT) to calculate the interannual Net XCO₂ variability over land for three latitudinal regions, Alaska/Canada, United States and the Amazon/Brazil. The OCO-2 design is optimized for sensitivity to XCO₂ variations, with an unprecedented combination of spatial resolution (about 3km) with narrow nadir coverage, while GOSAT provides broader spatial coverage (10km) with wider scanning coverage. There are different temporal degradations of both instruments over time because GOSAT was launched in 2009 and OCO-2 was launched in 2014. Both instruments infer CO₂ concentration from high-resolution measurements of reflected sunlight and use similar inversion algorithms to retrieve CO₂ concentrations. Both are passive satellites providing on-orbit global measurements of the greenhouse gas, XCO₂, for the years 2015 -2018. The results of the CS data fusion framework show that the fused data have Root Mean Square Error (RMSE) varying from 1.31 ppm to 4.12 ppm compared with original data, depending on the region of study and gridding resolution. Validation of fused data compared with AmeriFlux station towers observations shows RMSE of 2.68 ppm.

Keywords: Compressive sensing, data fusion, remote sensing, XCO₂.

1. INTRODUCTION

There are an increasing number of explorer and venture type satellites and/or ISS missions, aircraft Cubesats. In recent years, fusing multi-instrument information has gained extensive significance because merged information can often yield better insights and the inclusion of more temporal and spatial coverage over wider geographic scales for assessing regional and global climate changes.

Data fusion is the process of integrating multiple data sources to produce more consistent, accurate, and useful information than that provided by any individual data source. Sensor fusion, in particular, leads to higher probability, more accurate interpretations of readings (example shows in figure 1). This is a key task in this paper, as multiple sensors at different spectral resolutions or active and passive sensors respond differently to target signatures as well as under different external conditions such as different weather and cloud conditions. Data fusion processes can be used to combine raw or slightly processed level 1 signal data (brightness temperatures) or processed level 2 parameters (temperatures). Additionally, radar or lidar data usually have higher spatial resolution than IR or microwave data enabling fusing techniques to blend the best of both. Thus, we find that data carefully fused can be more informative than the original inputs.

* phuong3@umbc.edu 410-455-3535

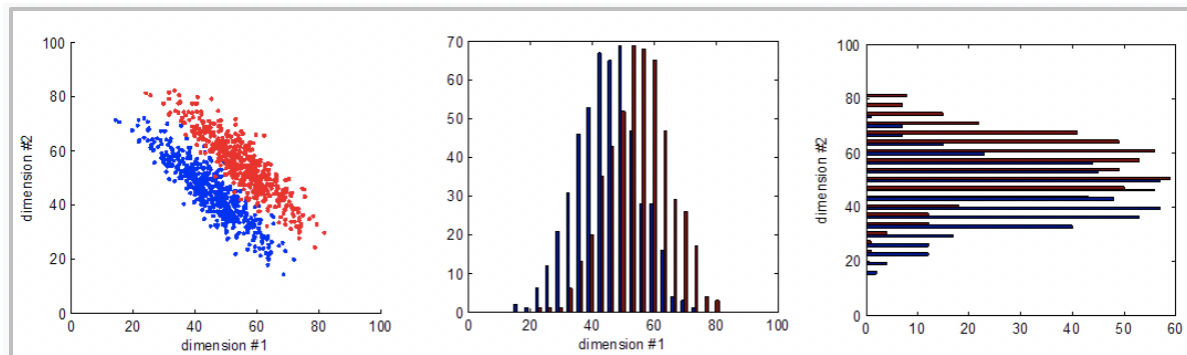


Figure 1 fusion of sensor readings from two sources (dimension #1 and dimension #2, center and right) can yield a superior (left) compared to classifiers based on a single sensor

The data fusion has been applied to ecological data examination [2]. Uptake of CO₂ over land using data fusion techniques is a great boon to the remote sensing experts, Earth scientists and researchers [2], as they could infer more accurate information on the carbon cycle from fusion of satellite data. Incorporating or combining accessible information from different sensors creates information on chemical and transport quantities that are not readily perceivable from a solitary sensor. In the field of remote sensing, daily satellite CO₂ data from the OCO-2 sensor provides narrow area coverage but fusing data from multiple satellites can provide broader spatial and temporal coverage. However, a major drawback to implementing data fusion techniques in remote sensing arises in needing to deal with inverse problems. Inverse problems arise when we have only a partial sample of measurements in the signal and need to reconstruct the full data set [10].

In addition, the challenges of fusing data from multiple sources include spectral, spatial, and temporal variations from different satellite imagery. For example, OCO-2 provides high spatial resolution and GOSAT gives wide spectral coverage measuring the total column-averaged dry-air mole fraction of CO₂ (XCO₂). OCO₂ overpasses the same location by 16 days while GOSAT gives a 3-day repeated cycle [21]. In addition, GOSAT observations decline about 2-4 ppm compared with OCO₂ and ground stations over time [25].

Compressive Sensing (also known as compressed sensing, compressive sampling or sparse sampling) is a recent sensing approach that outperforms the Nyquist–Shannon sampling frequency in acquiring and reconstructing sparse and compressible signals [13, 14, 15]. From an application point of view, compressive sensing has demonstrated outstanding performance where: (1) we are restricted by the factor of energy consumption on sensing side (e.g., wireless sensor networks, deep space missions); (2) we are limited to use few sensors or spectral channels (i.e., non-visible wavelengths, hyperspectral); (3) sensing is time-consuming (like medical imaging); or (4) measurement/sensing is too expensive (e.g., high-speed ADC), (5) where sensors are inhibited by data gaps (i.e. clouds, spatial coverage). From a signal processing perspective, Compressive Sensing exploits the sparsity of signals through optimization methods to reconstruct the original signal from far fewer samples than the imposed rate by the sampling theorem [16]. Compressive Sensing (CS) aims to recover a sparse signal. Utilizing Compressive Sensing (CS) techniques can help in the production of information, for example, in the estimation of information in areas with scant coverage. Applications of compressive sensing are generally acknowledged in the fields of clinical medicine, facial recognition, photography, cryptography, etc. Considering the positive aspects of CS theory in signal processing, we choose to use CS in studying the data fusion of CO₂ application. The question that the CS paradigm poses and solves is the ability to acquire signals from partial measurement coverage, and using convex optimization to reconstruct them over certain terrains. Even though the signals are acquired for regions with seemingly too few measurements, exploiting sparsity enables one to solve the resulting underdetermined systems of linear equations to efficiently recover the original signal [20]. Compressive sensing is a flourishing region of research with yet many open issues, such as non-asymptotic hypothesis of irregular lattices, geometric extrapolation, numerical analysis for various science domains. [1, 2, 6, 7]. Considering the positive aspects of CS theory in signal processing we choose to use this method in our problem.

The question that the Compressive sensing paradigm poses and solves is that acquiring signals with very few measurements, and reconstructing them exactly under certain conditions, using convex optimization. And even though the signals are acquired with seemingly too few measurements, exploiting sparsity enables one to solve the resulting underdetermined systems of linear equations to recover the original signal back, efficiently [20]. Compressive sensing establishments have joined with numerous regions of science like applied consonant investigation, outline hypothesis, geometric useful examination, numerical direct variable based math, enhancement hypothesis, and non-asymptotic hypothesis of irregular lattices [1, 2, 6, 7].

In this paper we propose to utilize Compressive Sensing (CS) method for data fusion. The evaluation of CS method will be performed by data fusion for the problem of the column-averaged dry-air mole fraction of CO₂ (XCO₂) estimation from Orbiting Carbon Observatory-2 (OCO₂) and Greenhouse Gases Observing Satellite (GOSAT) with additional validation from ground based stations (AmeriFlux). We choose two different Satellite observing XCO₂ systems OCO-2 and GOSAT to validate our proposed methodology. Since CO₂ is one of the most important greenhouse gases. Its concentration and distribution in the atmosphere have always been important in studying the Carbon Cycle and the Greenhouse effect. We will demonstrate that the fused data observations can complement each other in time and space and enrich the distribution data of XCO₂ globally. Then the fused data output will be validated by comparing with CO₂ observations from ground-based instruments of AmeriFlux towers.

2. RELATED WORK

Researchers are confronting difficulties with tremendous measures of information from sources that should be examined and broken down all around. Data fusion is one of the fundamental procedures in coordinating data from different sources to either accomplish refined or improved data of the spatial and additionally otherworldly perceptions [1, 2, 3]. Processing data from complementary observations from multiple instruments on the same platform, such as the (Environmental Office Solutions) EOS missions or even instruments from multiple satellites in orbits such as the A-train, can yield enhanced data products for studying the interactions of physical processes. If the multiple sensor data are measuring the same physical phenomenon such as temperature, then classical estimation methods such as Kalman filtering [4, 5] can be used to fuse the data. Fusing sensor information after each has obtained the observation's location, attributes, and identity. This can be solved using data fusion methods including weighted decision methods (voting schemas), inferences (Bayesian), and Demster-Shafers's method [6, 17]. The current and future multitude of sensors in less structured patterns of observations will need special tools (image registration) to register and produce validated enhanced sensor products [7].

Traditional data fusion focuses more on combining high-resolution panchromatic bands and low-resolution multispectral bands from the same sensor and both data sets are acquired at the same time. Those data fusion technologies include Brovey transformation, the intensity-hue-saturation (IHS) transformation, principal component substitution, wavelet decomposition, and Gram-Schmidt spectral sharpening method [8, 13]. Many other spatiotemporal data fusion methods have been developed to produce synthesized images with both high spatial and temporal resolutions from two types of satellite images, frequent coarse-resolution images, and sparse fine-resolution images [9]. Existing spatiotemporal data fusion methods are categorized into five groups based on the specific techniques used to link coarse and fine images: unmixing-based, weight function-based, Bayesian-based, learning-based, dictionary based, and hybrid methods [9]. Traditional and spatial temporal methods were designed based on different principles and strategies, and therefore show different strengths and limitations. The major difference between traditional data fusion and spatiotemporal data fusion is that traditional data fusion can not improve the temporal frequency of the original images.

Data fusion using pan-sharpening has been proposed to deal with the limitation of sensing observations in high resolution multispectral images using the combination of high spectral/low resolution and high resolution/narrow wavelength images [15]. The Principle Component Analysis (PCA) is applied to fuse the Landsat-TM multispectral and Spot Pan panchromatic images, achieving a good result [11]. PCA combined with wavelet transform is used to fuse images at the pixel level [16]. In addition, data fusion using cokriging interpolation has been applied for pan-sharpening multispectral Landsat ETM bands and showed the increasing spatial resolution [12].

Orthogonal Matching Pursuit (Compressive sampling) algorithm is used to fuse low resolution multiple spectral images and high resolution panchromatic images. High resolution image is produced by convolving the sparse coefficients (generated by correlating low resolution image patches with low resolution panchromatic) and high resolution panchromatic dictionary [14].

In [23], a procedure to fuse satellite images by methods for Compressive sampling matching Interest (CoSaMP) technique is proposed. This work produces inadequate coefficients by associating the low-resolution multispectral image with the low-goals skilnet word reference. Be that as it may, the nature of the reconstructed image isn't good. Image combination conspire for remote-sensing images is introduced by CS in the molded space in [24]. At first, the contourlet change is applied and the compressive samplings are intertwined by direct weighting. At long last, the image reproduction is completed by Iterative Edge Projection (ITP). Notwithstanding, this work is intended for remote-sensing images.

3. DATASETS

Satellite passive and active radiometers have been observing or inferring ocean and land surface data variables (i.e. temperature, albedo, vegetation, soil moisture, snow depth, topography, XCO₂, forest canopies, and chemical compositions such as CH₄, etc.) for decades now from different bands in the electro-magnetic spectral or pulsed Lidar reflectance signals. These sensor surface observations (e.g. ocean winds, temperatures, latent or sensible heat, solar radiation, etc.) when fused and assimilated into climate process proposed methods can be used to infer surface energy exchanges for monitoring global or regional temporal records of climate process changes, such as the Hydrology, Carbon or Biogeochemical cycles. The specific multiple Satellite Instrument datasets that we will focus on for this paper are:

Orbiting Carbon Observatory-2 (OCO-2) launched in July 2014 on a polar orbiting satellite and still producing operational XCO₂ and SIF level 2 accessible archived products are currently available from Sept. 1 through the present

Greenhouse gasses Observation Satellite “IBUKI” (GOSAT) was successfully launched Jan 23, 2009. The missions of GOSAT are to continue and enhance spaceborne measurement of major greenhouse gases (CO₂ and methane) from space GOSAT datasets are available from 2009 to present.

OCO-2 (provide high spatial resolution) and GOSAT (wide spectral coverage) are passive satellites on-orbit greenhouse gas satellites measuring global XCO₂ distribution. These two instruments have different observation strategies (resolution, spectral, orbits and temporal coverage). OCO₂ overpasses the same location by 16 days, GOSAT gives a 3-day period repeated cycle. Table 1 shows the brief description of these two satellites. Both instruments infer CO₂ concentration from high-resolution measurements of reflected sunlight and use similar inversion algorithms to retrieve CO₂ concentrations.

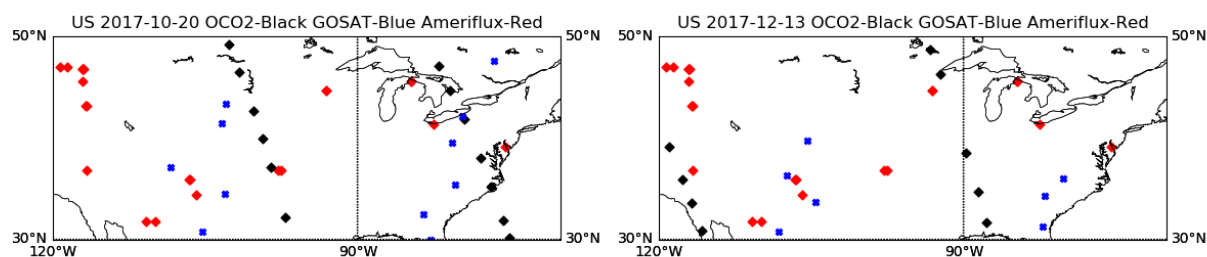


Figure 2a left and 2b right show the scan paths of OCO₂-black points and GOSAT-blue, and Ameriflux station-red points on two sample days. Observation's date and time ranges in a 2-3-hour time window.

In this paper we are concentrating on 3 geographical sites. The sites are (i) a region extending from Alaska and upper portion of western Canada at the same latitudes, (ii) continental United States and (iii) the Amazon basin consisting mostly of Brazil and Bolivia.

The data considered as input from this proposed method has been taken from two different satellite platforms. One is from the Orbiting Carbon Observatory (OCO-2) and the other is Greenhouse Gases Observing Satellite (GOSAT).

Both OCO2 and GOSAT sample the earth rather than filling the map. XCO2 from OCO2 provides high spatial resolution and GOSAT gives wide spectral coverage. Both OCO2 and GOSAT infer CO2 concentration from measurements of reflected sunlight and use a similar retrieval algorithm at SWIR (Short Wave Infrared Radiation). Level 2 datasets of XCO2 from OCO2 and GOSAT have been used.

We have used XCO2 (column-averaged dry-air mole fractions of CO₂). We have selected 3 regions for this study: Brazil/Amazon (Latitude -20 to 0, Longitude -70 to -40), the US (North) (Latitude 30 to 50, Longitude -120 to -70) and Alaska/Canada (Latitude 60 to 70, Longitude -165 to -90).

| | GOSAT | OCO2 |
|-----------------------|--|--|
| Orbit high | 675 km | 705 km |
| Spatial Resolution | 10.5 km | 1.29 km x 2.25 km |
| Observation time | 3 day repeated cycle Total 44 orbits 98.1 min | 16 day repeat cycle Total 233 orbits 98.8 min |
| Equator crossing time | 1:00 pm | 1:36 pm |
| Spectral frequency | 0.76, 1.6 and 2.0 μm | 1.61 and 2.06 μm |
| Launch date | Jan 2009 | July 2014 |

Table 1 OCO2 and GOSAT brief description

We considered a rectangular grid for each of the sites by considering the latitude and longitude. From the OCO-2 and GOSAT satellite we have chosen time, latitude, longitude, XCO₂ concentration observation for our proposed method. For each site we applied the latitude and longitude in the form rectangular grid and collected all the XCO₂ concentration observations present inside the rectangular grid.

The observation data obtained from each site has data observed for every second for four years due to which the data is huge for each site. The input considered for this proposed method has data for four years which is averaged on a daily basis as the data observed is for every second. So, we average the observations present for each day for all three sites. This is done for both the satellite stations GOSAT and OCO-2.

4. METHODOLOGY

We have developed a CS data framework and tested it to fuse CO2 observations from two different satellites (low and high resolution), OCO2 and GOSAT to produce high quality gridded datasets. We have validated the method using data coverage of three different regions (over US, Alaska, Brazil/Amazon) and compared with ground based stations dataset (observations of CO2) over the US region.

Compressive Sensing approaches emphasize the use of constrained optimization methods of underdetermined linear systems along with spectral methods. These techniques typically assume that the scene can be represented as a sparse linear combination of basis functions. A popular set of basis functions for this purpose are the 2D wavelets, and notable compression factors of 5x SNR have been achieved using the LASSO regularization method as follows,

$$\min ||Ax - b||_2^2 + \lambda ||x||_1$$

LASSO regularization attempts to find the solution of $Ax = b$ such that the L1 norm is less than a given constant and is reformulated as an unconstrained optimization problem by the method of Lagrange multipliers. LASSO optimization is convex and reducible to linear programming, and thus solvable in polynomial time by interior point methods [26, 27].

In practice, fast implementations have been achieved using Bregman iteration [28]. The L1 norm minimization problem is an approximation to the L0 norm minimization for sparse signal reconstruction, and many alternative techniques exist including greedy algorithms such as Matching Pursuit (MP) and its derivatives. There are also methods that attempt to reweight the L1 regularization problem into a space that guarantees equivalence with the L0 regularization norm. A widely studied extension to compressive sensing is the notion of robust statistics using M-estimators which are resistant to outliers. A review of robust compressive sensing algorithms using a variety of M-estimators can be found by [29]. Three common M-estimators include the L2 loss, the L1 loss, and the Huber loss norms for the $Ax - b$ term all lead to convex optimization methods. Both the L1 and Huber norms have been shown to improve resilience and reduce breakdown when applied to compressive sensing with outliers in observation vector b . In a comparison of M-estimator techniques, [29] found the L1 formulation of to exhibit a very good compromise between performance and resilience to outliers in measurement vector b as follows,

$$\min ||Ax - b||_1 + \lambda ||x||_1$$

From a signal processing perspective, Compressive Sensing exploits the sparsity of signals through optimization methods to reconstruct the original signal from far fewer samples than the imposed rate by the sampling theorem. We assume that combination of GOSAT and OCO2 observations have both spatial and temporal correlations. Let $\mathbf{x} \in \mathbb{R}^n$ represent spatial-temporal observation data collected by GOSAT and OCO2 instruments in the sensing period from 2015 to 2018 period. Where i th element of \mathbf{x} is either GOSAT or OCO2 observations (depending its observed time t). CS has been used for the reconstruction of the \mathbf{x} signal (all observations) given the formulated \mathbf{y} (compressive measurements). Matrix A is created using the inverse discrete cosine transform of an identity matrix. L1-norm optimization problem is used to create reconstructed high probability signals with its inverse cosine transform back to its original signal space. Both Root Mean Square Error (RMSE) and R2 correlation coefficient have been used to compare between reconstructed signals and the original signal. Then the reconstructed signals are compared with AmeriFlux ground station data for validation.

5. EXPERIMENTS AND EVALUATION

We performed fusion dependent on CS as portrayed in the past segment. This session will show the evaluation of the results. This evaluation involves a comparison with the observation measurement using the distributed ground based station data (AmeriFlux towers). The error and correlation are used as evaluation metrics. For now, we evaluate our fusion methods through a few quantitative methods. We evaluate both the temporal and spatial quality of our fused data. Below sessions will show the evaluation of our methods at different resolutions (5.1), comparison with ground station datasets (5.2), and performing at different compression ratio (5.3).

5.1 Evaluation at different resolutions and analysis

So as to assess the spatial nature of the combined data we figure the gridding of the OCO2, GOSAT independently and OCO2 and GOSAT fused datasets. Then we perform the comparison. The thought is that the combined dataset ought to have upgraded data content contrasted with their comparing input OCO2 and GOSAT independently. Therefore, the lower the errors of the fused information is contrasted with its relating either input data, the better the spatial nature of the combination is. We conduct two types of experiments. One with fine and other with coarse grid regions.

5.1.1 Evaluation at the high grid resolution

We aggregate (average) OCO2 observations by 30 seconds for reducing processing time and variations in OCO2 observations at high resolution. This method can be considered as gridding high resolution OCO2 data at approximately 3x2.5 degree. The gaps in data observations of both OCO2 and GOSAT are preserved.

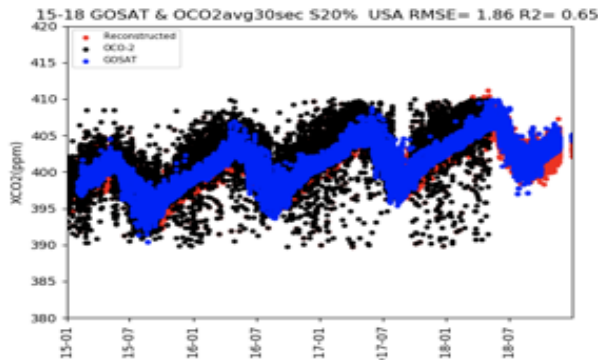


Figure 3a USA region 2015-2018

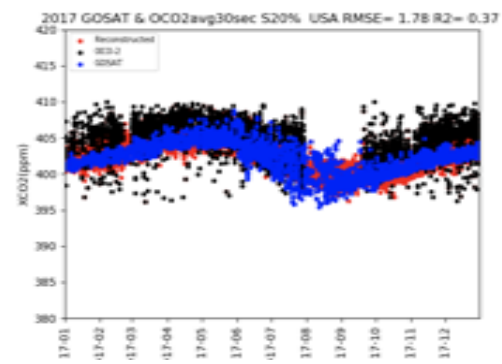


Figure 3b USA year 2017

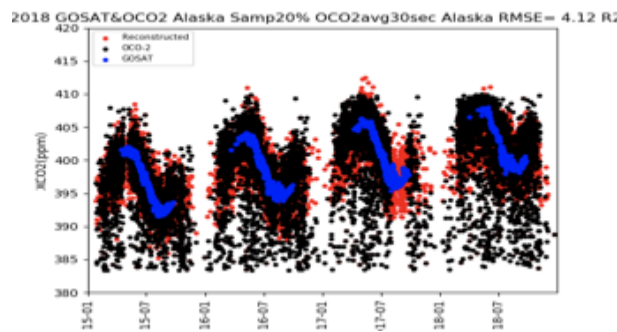


Figure 4a Alaska region 2015-2018.

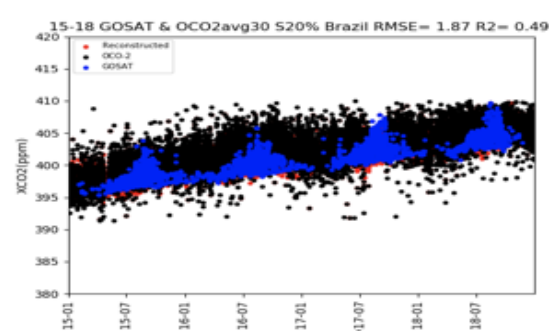


Figure 4b Brazil (Amazon) region

Figure 3a and 3b shows the results of reconstructed data (fused data) points (red) compared to the original signal of GOSAT (blue) and OCO2 (black). Figure 3a also shows that there are periods in which we do not have OCO2 (2018 from July to the end of the year), the framework still produces fused output data. Figure 3b shows the gaps from the middle of July to early oct 2017 where OCO2 does not have observation datasets. The framework produced estimated signals which have RMSE of 1.78 ppm and $R^2=0.37$ compared to the original signal from mixed GOSAT and OCO2. OCO2 has a high resolution of 3km so one can see there are many points within the 15km observed region of GOSAT as well. Figure 4a and 4b plot results of reconstructed data (fused data) for 4 years over Alaska and Brazil respectively also at 3x2.5 degree resolution. As seen in figure 4a there is more variation in OCO2 observations because of the region's latitudinal coverage and it impacts the fused signal with higher errors of 4.12 ppm and low correlation 0.08. Over the Brazil region the RMSE error is lower with 1.87 ppm and fused signals correlated better with original signal from OCO2 and GOSAT ($R^2=0.49$).

5.1.2 Evaluation at lower resolution daily, monthly over three regions

The second set of experiments (figure 5a, 5b, 5c), we grid OCO2 and GOSAT separately at big regions and produce daily dataset before we merge them together for this data fusion technique.

This gridding method is equivalent with gridding at 30x50 degree over US, 10x75 degree over Alaska , and 20x30 over Brazil regions. These resolutions are the order of magnitude lower than that of resolution used in experiments shown in session 5.1.1.

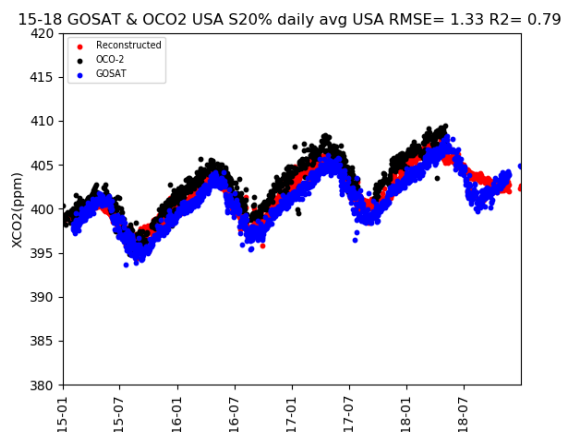


Figure 5a sample 20% average both GOSAT and OCO2 OCO2 by day

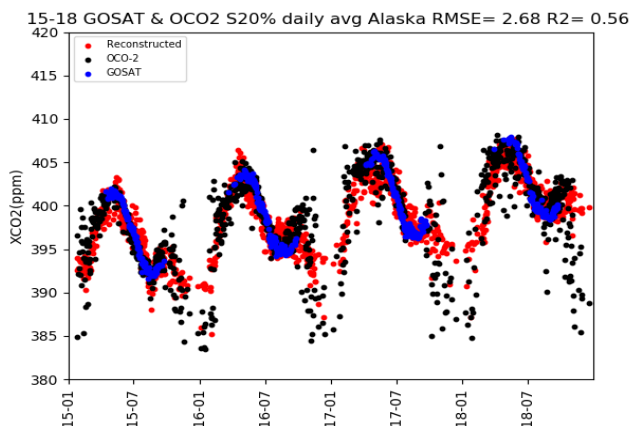


Figure 5b sample 20%, average both GOSAT and by day

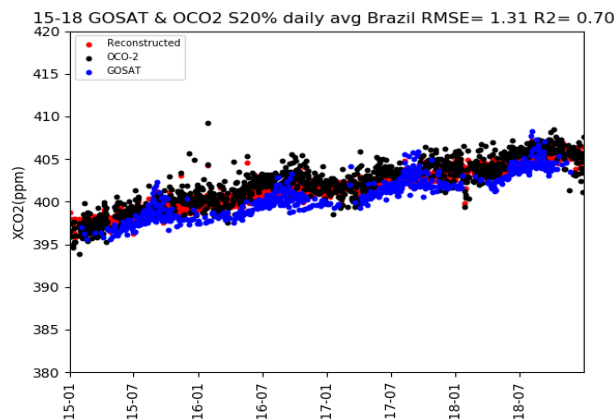


Figure 5c daily avg Barzil with 20% samples

The constructed XCO₂ values from the CS algorithm are complemented by OCO₂ observations via versus with lower errors and higher correlation in all three regions compared with previous results when we used 3x2.5 resolution (higher resolution).

Figure 6a, 6b and 6c show the monthly average of fused data over three regions when running experiments using daily average data in figure 5a, 5b, and 5c. As shown in the monthly average XCO₂ of OCO₂ and GOSAT satellites (figure 6a, 6b and 6c), and our reconstructed XCO₂ for the year 2015 for Brazil/Amazon, the US (North) and Alaska/Canada respectively.

Monthly mean of GOSAT & OCO2 data fusion United States, year 2015

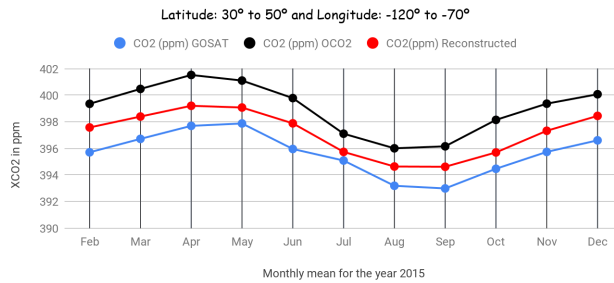


Figure 6a United State 2015: OCO2 and reconstructed OCO2/GOSAT RMSE=1.89 R2=0.97. GOSAT and reconstructed OCO2/GOSAT RMSE=1.55 R2=0.95

Monthly mean of GOSAT & OCO2 data fusion Alaska, year 2015

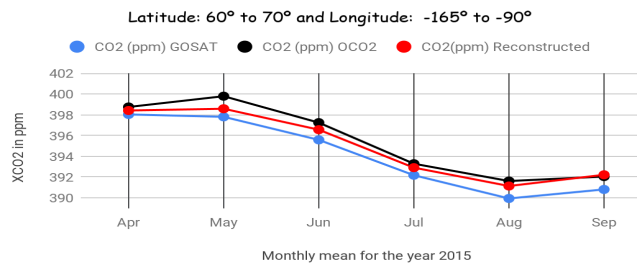


Figure 6b Alaska, 2015: OCO2 and reconstructed OCO2/GOSAT RMSE=0.64 R2=0.98. GOSAT and reconstructed OCO2/GOSAT RMSE=0.97 R2=0.99

Monthly mean of GOSAT & OCO2 data fusion Amazon/Brazil, year 2015

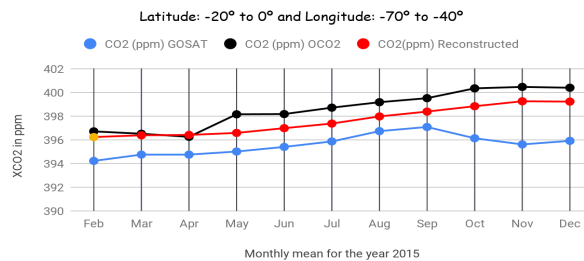


Figure 6c Amazon, Brazil 2015: OCO2 and reconstructed OCO2/GOSAT RMSE=1.2 R2=0.91. GOSAT and reconstructed OCO2/GOSAT RMSE=2.16 R2=0.51

The results in figure 6a show that GOSAT observations has decreased and lagged behind OCO2 (about 1.5 to 4 ppm) in the ability to monitor both North and South America areas. Ailin Liang, et al also found a similar decline in GOSAT compared with OCO2 and ground stations. Provided above experiments are the results of an increase in accuracy of GOSAT satellite data after application of algorithm using OCO-2 satellite data regarding Carbon Dioxide in the atmosphere using our Compressive Sensing data fusion algorithms. Our results showed that reconstructed XCO2 are closer to OCO2 observation where the GOSAT drifted further down.

5.2 Performance analysis again ground based tower dataset

Quality appraisal of fused data is fundamental to assess the potential advantages of data fusion strategy just as to contrast the outcomes obtained and various algorithms. For validation of fused datasets, there are two main approaches. One can use different methods to produce datasets and compare. Other people can use different sources of datasets which measure the same physical phenomenon for comparison. In this section we use AmeriFlux station dataset (the later approach) as they measure the CO₂ quantity. The AmeriFlux towers have either hourly or half an hourly observations. Thus we match our reconstructed data point with AmeriFlux data point by hour, half hour time period. Then, the average by latitude band is performed on both AmeriFlux and our reconstructed data from using CS method before comparison. Since there are less number of AmeriFlux stations and the satellite cover paths and observation points.

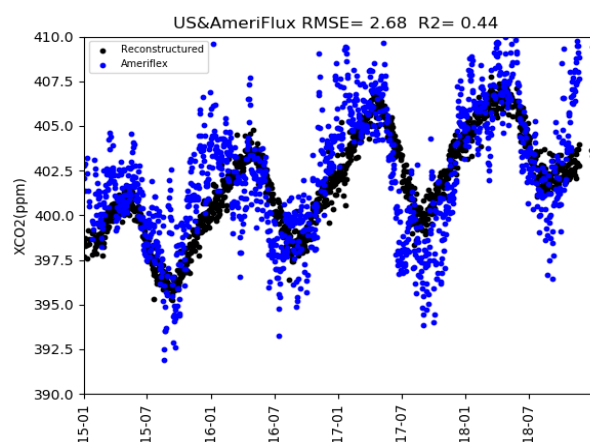


Figure 7a: Using gridding 30 seconds average results from

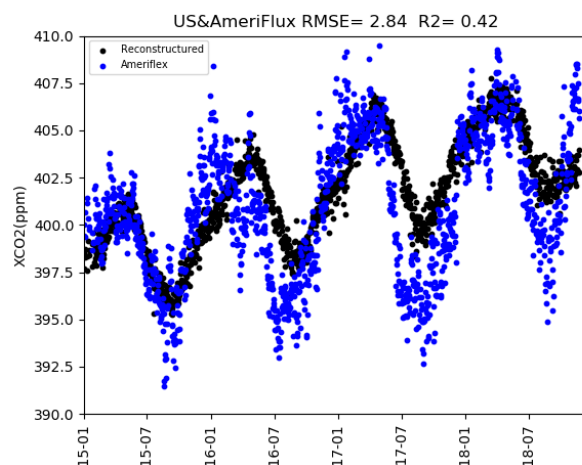


Figure 7b Using daily average results from

Over US. The data is average over the US lat: 30N-50N, lon 120W-70W before comparison (20 latitude band). The results in figure 7a and 7b show that the fused data output from using our method has errors of 2.68 ppm and 2.84 ppm over the US region. Within the US region, we only can process 23 AmeriFlux stations. Thus it limits the coverage of station datasets for comparison. However, this error rate is less than the GOSAT instrument drifted errors and within bias validation between instrument observation compared with ground station observations [25]. Notice that we used only 20% of samples (compressed signal) to recover the 100% signal in the experiments. The RMSE errors will be lower when the compressed ratio increases (see session 5.3).

5.3 Performance analysis at different compressed ratio

The three regions considered for our experiment are evaluated on the basis of errors and correlation varied with the percentage of samples for each region shown in Fig.8. We observe that the reconstructed data error for the US has decreased from 1.04 ppm to 0.65 ppm. This indicates that more the number of samples less are the errors. In the same way when we consider the correlation for the US region it is the inverse as we see in the below Figure 8. as it shows better correlation when the number of samples is more.

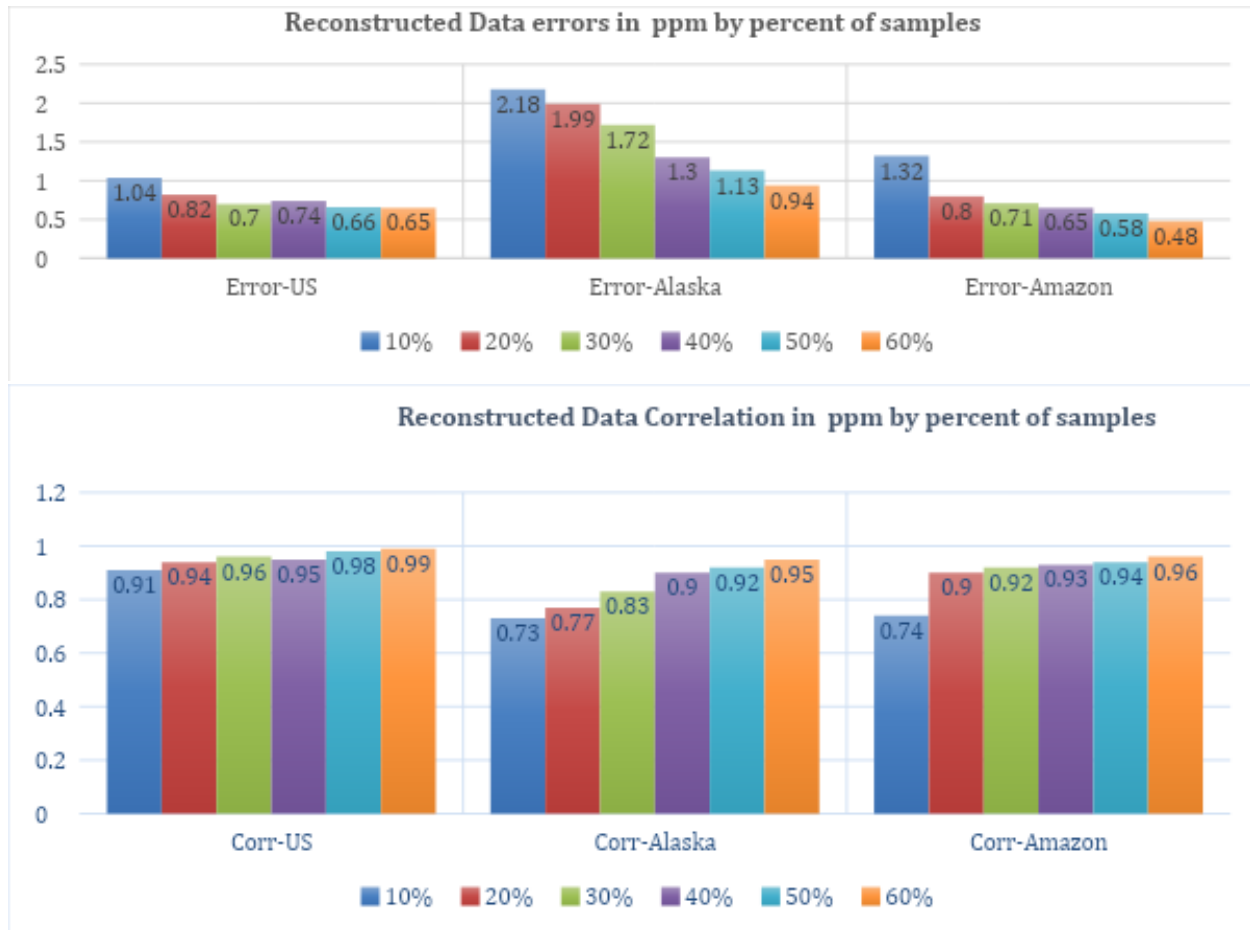


Figure 8. Comparison for error and correlation in ppm for all three regions based on percent of samples

6. CONCLUSION

We have developed a satellite sensor-based Data Fusion framework employing a Compressive Sensing method. The goals are to achieve refined or improved observations over time and space. The framework has been used to test and validate data fusion of XCO₂ from multiple satellites namely OCO₂ and GOSAT. We have shown that the fused data outputs can complement both satellite instruments' observations in time and space. Hence, it can enrich the distribution data of XCO₂ globally. Our evaluation results show that the fused data have RMSE varying from 1.31 ppm to 4.34 ppm (part per million) compared with the original data. Data validation by comparison with AmeriFlux station tower data shows RMSE of 2.68 ppm over US regions. Our proposed sensor data fusion method can be used to generate the estimation of XCO₂ datasets for exploring and analyzing the Global Carbon Cycle.

ACKNOWLEDGMENTS

This study was funded by the NASA grant number NNN16ZDA001N-AIST16-0091. We also acknowledge the support of the NSF supported Center for Accelerated Real Time Analytics (CARTA), University Of Maryland Baltimore County

REFERENCES

- [1] Gunho Sohn, Ian Dowman, Data fusion of high-resolution satellite imagery and LiDAR data for automatic building extraction, *ISPRS Journal of Photogrammetry and Remote Sensing*, Volume 62, Issue 1, 2007, Pages 43-63, ISSN 0924-2716
- [2] Stefan R. Falke, Rudolf B. Husar & Bret A. Schichtel (2001) Fusion of SeaWiFS and TOMS Satellite Data with Surface Observations and Topographic Data during Extreme Aerosol Events, *Journal of the Air & Waste Management Association*, 51:11, 1579-1585, DOI: 10.1080/10473289.2001.10464386
- [3] Liggnis, M.E.; Hall, D.L.; Llinas, J. *Handbook of Multisensor Data Fusion*, 2nd ed.; CRC Press: Boca Raton, FL, USA, 2009; eBook ISBN 9781315219486
- [4] Kalman, R. E. 1960. "A New Approach to Linear Filtering and Prediction Problems," *Transaction of the ASME—Journal of Basic Engineering*, pp. 35-45 (March 1960).
- [5] G. Welch and G. Bishop An Introduction to the Kalman Filter Available from <http://www.cs.unc.edu/welch/kalman/kalmanIntro.html>
- [6] Huadong Wu, M. Siegel, R. Stiefelbogen and Jie Yang, "Sensor fusion using Dempster-Shafer theory [for context-aware HCI]," *IMTC/2002. Proceedings of the 19th IEEE Instrumentation and Measurement Technology Conference (IEEE Cat. No.00CH37276)*, Anchorage, AK, USA, 2002, pp. 7-12 vol.1. doi: 10.1109/IMTC.2002.1006807
- [7] J. Le Moigne, N. Netanyahu, & R. Eastman (Eds.), *Image Registration for Remote Sensing* (pp. I-Vi). Cambridge: Cambridge University Press.
- [8] Zhu, X.; Chen, J.; Gao, F.; Chen, X.; Masek, J.G. An enhanced spatial and temporal adaptive reflectance fusion proposed method for complex heterogeneous regions. *Remote Sens. Environ.* 2010, 114, 2610–2623
- [9] Xiaolin Zhu, Fangyi Cai, Jiaqi Tian and Trecia Kay-Ann Williams. "Spatiotemporal Fusion of Multisource Remote Sensing Data: a Literature Survey, Taxonomy, Principles, Applications, and Future Directions"
- [10] D. L. Hall, *Mathematical techniques in multisensor data fusion*. Norwood: Artech House Inc, 1992.
- [11] P. S. Chavez, S. C. Sides, and J. A. Anderson, "Comparison of three different methods to merge multiresolution and multispectral data: Landsat TM and SPOT panchromatic," *Photogrammetric Engineering and Remote Sensing*, vol. 57, no. 3, pp. 295–303, March 1991
- [12] N. Memarsadeghi, J. Le Moigne and D. Mount, "Image Fusion Using Cokriging," 2006 IEEE International Symposium on Geoscience and Remote Sensing, Denver, CO, 2006, pp. 2518-2521. doi: 10.1109/IGARSS.2006.651
- [13] J. Nunez, X. Otazu, O. Fors, A. Prades, V. Pala and R. Arbiol, "Multiresolution-based image fusion with additive wavelet decomposition," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 3, pp. 1204-1211, May 1999. doi: 10.1109/36.763274
- [14] B. Sathyabama, S. G. Siva Sankari and S. Nayagara, "Fusion of satellite images using Compressive Sampling Matching Pursuit (CoSaMP) method," 2013 Fourth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), Jodhpur, 2013, pp. 1-4. doi: 10.1109/NCVPRIPG.2013.6776256
- [15] Y. Zhang, "Understanding image fusion," *Photogrammetric Engineering and Remote Sensing*, pp. 657–661, June 2004.

- [16] V.P.S. Naidu and J.R. Raol, "Pixel-level Image Fusion using Wavelets and Principal Component Analysis". Defence Science Journal, Vol. 58, No. 3, May 2008, pp. 338-352 Ó 2008, DESIDOC.
- [17] Z. Hongwei , O. Basir , F. Karray , Data fusion for pattern classification via the dempstershafer evidence theory, in: IEEE International Conference on Sys- tems, Man and Cybernetics, 2002, pp. 109–110
- [18] Wei, J.; Wang, L.; Liu, P.; Chen, X.; Li, W.; Zomaya, A.Y. Spatiotemporal Fusion of MODIS and Landsat-7 Reflectance Images via Compressed Sensing. IEEE Trans. Geosci. Remote Sens. 2017, 55, 7126–7139.
- [19] - Donoho, David L. "Compressed sensing." IEEE Transactions on information theory 52.4 (2006): 1289-1306.
- [20] - Foucart, Simon, and Holger Rauhut. "An Invitation to Compressive Sensing." A mathematical introduction to compressive sensing. Birkhäuser, New York, NY, 2013. 1-39.
- [21] - P. Nguyen, M. Halem, J.Dorband, A. Radov, D. Frankel, “ “Evaluation of Deep Learning proposed methods for Predicting CO2 Fluxes”, AGU Fall Meeting, Dec. 2017, New Orleans
- [22] - He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
- [23] B. Sathyabama ; S.G. Siva Sankari ; S. Nayagara, "Fusion of satellite images using Compressive Sampling Matching Pursuit (CoSaMP) method", Fourth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), 18-21 Dec, Jodhpur, India, 2013.
- [24] Yang Senlin ; Li Yuanyuan ; Wan Guobin, "Remote-sensing images fusion by compressed sensing in contourlet transform domain", IEEE Workshop on Advanced Research and Technology in Industry Applications (WARTIA), Ottawa, ON, Canada, 29-30 Sep, 2014
- [25] Ailin Liang, Wei Gong, Ge Han and Chengzhi Xiang, “Comparison of Satellite-Observed XCO2 from GOSAT, OCO-2, and Ground-Based TCCON”, 2017.
- [26] Candes E. J., & Tao, T. (2005) “Decoding by linear programming,” IEEE Transactions Information Theory, vol. 51, no. 12, (pp. 4203–4215).
- [27] Kim, S. J., Koh, K., Lustig, M., Boyd, S., & Gorinevsky, D. (2007). An interior-point method for large-scale ℓ_1 -regularized least squares. *IEEE journal of selected topics in signal processing*, 1(4), 606-617.
- [28] Goldstein, T., & Osher, S. (2009). The split Bregman method for L1-regularized problems. *SIAM journal on imaging sciences*, 2(2), 323-343.
- [29] Carrillo, R. E., & Barner, K. E. (2013). Lorentzian iterative hard thresholding: Robust compressed sensing with prior information. *IEEE Transactions on Signal Processing*, 61(19), 4822-4833.