Mapping Titan's dunes and interdune regions with Machine Learning

L. E. Bonnefoy*

Planetary Sciences department, University of Arizona

(Dated: December 14, 2016)

Mapping Titan dunes can lead to important analyses of their morphologies, orientations, and extent. We developed a machine learning algorithm to map individual dunes and interdune regions. The best algorithm we find is a Support Vector Machine with an accuracy of 79.9%. While it inaccurately maps their width and spacing, it successively identifies dune fields and maps most dunes well.

I. INTRODUCTION

Titan, satellite of Saturn and second largest satellite in the Solar system after Ganymede, is in many ways very similar to Earth. Indeed, the 94K surface temperatures allow the presence of liquid methane and ethane on the surface, making Titan the only body in the Solar System other than Earth with free-standing large bodies of liquids on its surface. The Cassini-Huygens mission to the Saturn system has made many important discoveries on Titan, revealing lakes and seas near the poles (Stofan et al. (2007)), mountains (Radebaugh et al. (2007)), evidence for precipitation (Turtle et al. (2011)), fluvial networks (Burr et al. (2013)), possible cryovolcanism (Soderblom et al. (2009), Lopes et al. (2013)), and extensive dune fields (Lorenz et al. (2006),Radebaugh et al. (2008)). These linear dunes are confined within $\pm$ 30° of the equator, can extend for more than a hundred kilometers, and cover approximately 15% of Titan's surface (Rodriguez et al. (2014), Lopes et al. (2016)), making them the largest known surface reservoir of organics.

Titan's dunes can teach us a lot on Titan's history, atmospheric conditions, erosion, and sediment transport. The orientations of the dunes have been mapped and studied by various groups, who linked them either to long-term climate cycles due to Titan's orbit (Ewing et al. (2015), McDonald et al. (2016)), or to strong methane storms (Charnay et al. (2015), Lucas et al. (2014b)). The morphology of the dunes, i.e. their width, spacing, length, and shape, show latitudinal and elevation trends that reveal variations in sediment availability (Le Gall et al. (2012), Savage et al. (2014)). Finally, mapping dune fields have revealed that the dunes have a unique composition on Titan: they are the second darkest features in both infrared and Radar

*Electronic address: bonnefoy@lpl.arizona.edu

data after the methane/ethane lakes (Soderblom et al. (2007), Le Gall et al. (2011), Le Gall et al. (2014), Lopes et al. (2016)). Lucas et al. (2014a) developed a non-local denoising algorithm to clean the Radar data, before using a Linear Segment Detection algorithm to identify individual crestlines. Mapping individual dune area has been done once, to extract the separate radiometry and infrared signatures of the dune and interdune regions (Bonnefoy et al. (2016)). This study used a local thresholding algorithm to distinguish between dune and interdune areas within each dune field. Not only does this method perform poorly when the contrast between dune and interdune Radar brightness is small, it also will identify any uniformly dark sand sea as dunes, thus being applicable only within dune fields. Herein, we attempt to develop an automated dune mapping algorithm that corrects for these deficiencies.

While dune mapping machine learning algorithms have been developed on Mars (Bandeira et al. (2011)), they are not applicable to Titan. Indeed, the resolution (300 m/pixel), data type (Synthetic Aperture Radar images), and landscapes are too different. We can, however, apply similar machine learning algorithms to Titan. Here, we apply both Support Vector Machines (SVMs) and Neural Networks (NNs) to map individual dunes on Titan. In Section II, we describe the Cassini Synthetic Aperture Radar (SAR) dataset, our labeling method, and the optimization of the algorithm's parameters. In Section III, we describe the accuracy of the resulting algorithm and its strengths and weaknesses.

II. METHODS

A. Data description

The Cassini Radar operates at 2.16 cm (13.78 GHz) in order to penetrate Titan's thick absorbing atmosphere. In Synthetic Aperture Radar (SAR) mode, it can obtain images of Titan's surface at a resolution of $\approx$ 300 m/pixel (lower for high-SAR, obtained further from Titan). Since Titan's dunes are about 800 m to 2 km wide and are separated by 1 to 5 km interdune regions (Le Gall et al. (2012), Savage et al. (2014)), SAR is able to resolve them. The intensity of the SAR images is a function of composition (dielectric constant), surface slope, and surface structure (particle size, volume scattering). Consequently, the dunes, which are composed of sand-sized particles and have the second lowest dielectric constant on Titan, are usually dark to SAR. In a few occasions where the incidence angle was perpendicular to the slope of the dune, quasi-specular effects caused the dunes to appear bright in SAR (Radebaugh et al. (2010)); this data was ignored. The regions separating the dunes, called interdune regions, are brighter, most likely because they are sand-free and have a distinct composition (Callegari et al. (2014), Bonnefoy et al. (2016)).

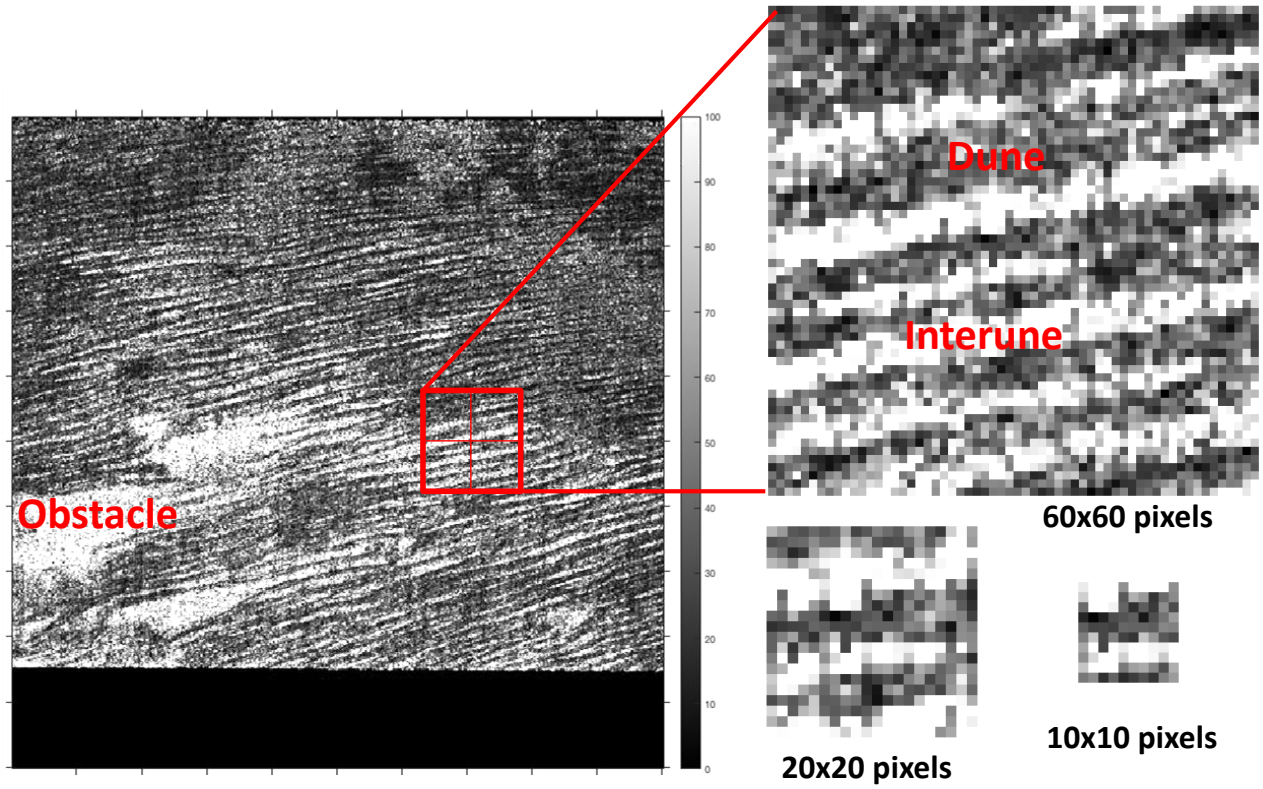

Figure 1: Illustration of the labeling method on a section of SAR swath t17. Left: a 500x500 pixel section of the SAR swath. When the user clicks on a point within this image, three enlarged views are shown, of 10x10, 20x20, and 60x60 pixels (right). The user then inputs whether the center point is a dune (1), an interdune region (2), or another terrain (0). This point would be labeled as dune.

We also note that, since the Cassini spacecraft is in orbit around Saturn rather than Titan, the SAR dataset is composed of different swaths, where each swath was taken during a distinct flyby of Titan.

B. Labeling

Before we can train and test an algorithm to recognize dunes, interdune regions, and other terrains, it is necessary to label a training dataset. We therefore developed a labeling code which displays 500 by 500 pixel sections of a swath. A user click on the image opens three enlarged views of the region, of 10x10, 20x20, and 60x60 pixels (see Fig. 1). The user then labels the image as dune (1), interdune region (2), or other (0), and the label and enlarged images are recored into a file. The label corresponds to the center point of the enlarged images, as no image will be exclusively dune or interdune region. In order to account for varying dune and interdune Radar brightness, dune orientation, and dune width, 1257 points were labeled on three different SAR swaths (t17, t55, and t61). As shown in Fig. 2, these labels are well-balanced, which is necessary for the calculated accuracies to be valid.

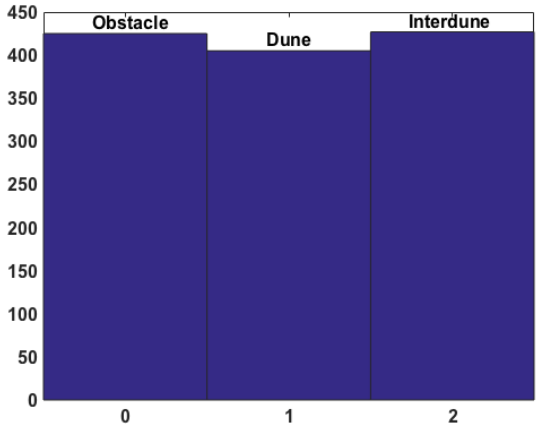

Figure 2: Distribution of the labeled data. There are a total of 1257 labeled points.

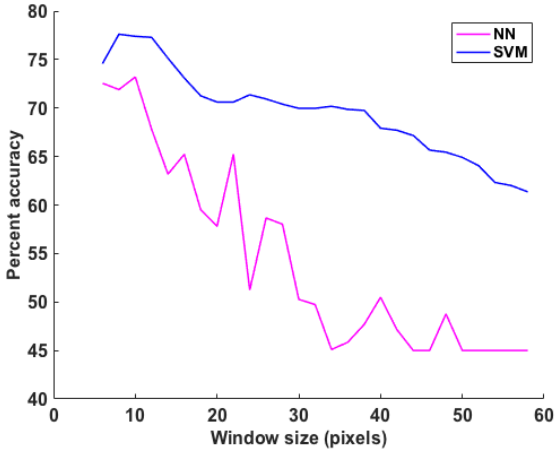

Figure 3: Accuracy of the neural network (NN) and support vector machine (SVM), using different window size. The SVM always performs significantly better than the NN.

C. Testing and optimizing

We have tested both Neural Networks (NN) and Support Vector Machines (SVM) on this dataset before finding the optimal option.

1. *Support Vector Machine*

The principle of Support Vector Machines is to find the optimal hyperplane that maximizes the margin between the different classes (Boser et al. (1992)). This is done linearly in a hyperspace defined by a kernel. In order to optimize the SVM parameters, we ran a grid search over three different kernels (linear, rbf, and polynomial) and over their respective parameters. For the linear and rbf kernels, we took $\gamma = 1E(-6\ to\ 0)$, and for the polynomial kernel we took degree = 2 to 4. We found that the accuracy was highest for the rbf (radial basis function) kernel with $\gamma = 1E-5$. The formula for the rbf kernel is:

$$K(x, x') = exp(-\gamma||x - x'||^2)$$

2. *Neural Network*

A perceptron inputs a number of weights and a bias, then uses an activation function to output a certain value. A neural net is composed of one or several connected layers of perceptrons leading to an output. The neural network is trained using feedforward-backpropagation: all

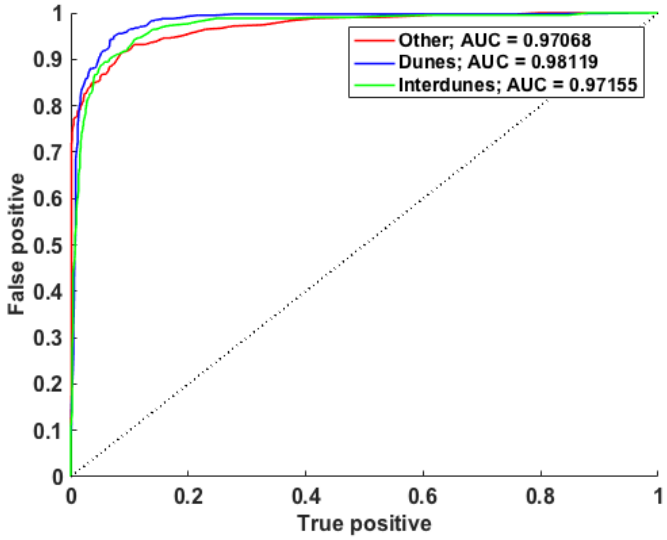

Figure 4: ROC plot for dunes, interdunes, and other terrains. The areas under each curve are given.

units are evaluated to lead to the label, the error on the label is evaluated, and the error is propagated back through the layers, adjusting the weights progressively. The parameters we needed to optimize for the neural network were the number and size of layers, the activation function, and the regularization parameter α. We also used a grid search to optimize these parameters. The best parameters varied as a function of the window size.

3. Optimizing window size

The last parameter to consider with this labeled dataset, for both SVMs and NNs, is the window size. Indeed, with a small window (Fig. 1, bottom right) the algorithm will more easily recognize the importance of the brightness contrast between dune and interdune. In a large window (Fig. 1, top right), the linear dune/interdune pattern will be more apparent. To find the best window size, we optimized both the neural network and the support vector machine for window sizes ranging from 6 to 58 pixels. As shown in Fig. 3, the SVM always performs better than the NN. We therefore used an SVM with an rbf kernel, $\gamma = 1E-5$, and a window size of 10x10 pixels.

III. RESULTS

Our final algorithm has an accuracy of 79.9%, as calculated using 5-fold cross-validation. We plot the ROC for each of our three labels, as shown in Fig. 4. The areas under the curves are high for all three labels and highest for the dunes, indicating that we have a good classifier.

In order to generate images using our classification algorithm, we iterate through the image

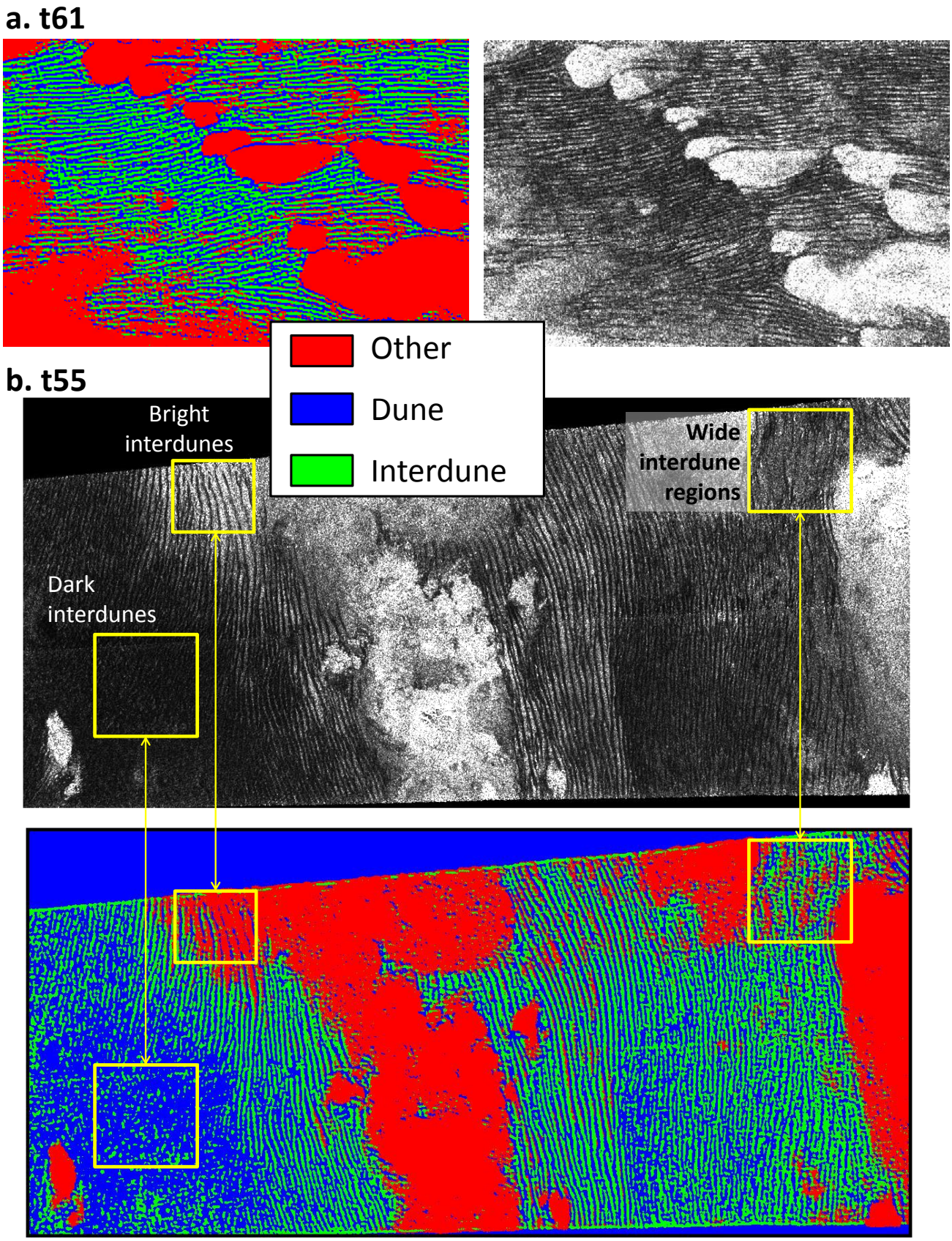

Figure 5: Extracted masks and original SAR data. a. (left): Classified section of SAR swath t61. (right) SAR data in that region. b (top): Classified section of SAR swath t55. (bottom) SAR data in that region.

with a (flattened) box of 10x10 pixels (the optimal window size). The label we obtain will correspond to the label of the center pixel. We can thus generate a map where each point is labeled as dune, interdune, or obstacle. We show in Fig. 5 two examples or such maps. Notice that the vertical dunes in t55 and the horizontal dunes in t61 are equally well mapped: the algorithm's performance is not hindered by dune direction.

Because of the small window size, the algorithm only partially sees the linear dune/interdune pattern. The consequence of this fact is, first, that wide interdune regions will be poorly mapped, as they will be interpreted as "other". The second consequence is that the algorithm will instead really heavily on simple thresholding methods. Indeed, dunes are always the darkest feature in an image, obstacles are generally the brightest, and the interdune regions lie somewhere in between. Therefore, dark regions, such as areas with dark interdunes or more simply the outside of the swath (it is equal to zero everywhere) will be mapped as dunes. On the reverse, very bright interdune regions will be mapped as "other". These problems are all shown in Fig. 5. Even within well-behaved regions, the algorithms tends to uniformize the dune width and spacing.

IV. CONCLUSION AND FUTURE WORK

The applicability of our maps is partly hampered by its inaccuracies. Indeed, only areas with small interdune regions, high dune/interdune contrast, and not too bright interdunes will be accurately mapped. Any calculation of dune or interdune area is not possible with these maps. Measuring dune width and spacing statistics is also inaccurate. It is feasible, however, to study dune orientation from our maps. Yet dune orientation is more simply mapped with a Linear Segment Detection algorithm, as done by Lucas et al. (2014b). Our algorithm is good, but needs improvement to be applicable.

There are several directions to go to refine the algorithm. The first would be to apply it to the denoised dataset created by Lucas et al. (2014a). The lack of noise should make recognizing features considerably easier, and should decrease the number of false positives/negatives. Another possible improvement would be to change the labeling method. Instead of labeling points, it is possible to map some dunes and interdunes by hand. We do note that, while this method would avoid some labeling bias due to the ambiguity of the dune edge, it would be much harder to have a complete dataset, that takes into account variations in dune direction, dune and interdune brightness, and dune morphology. It is also possible to apply a convolutional neural network to this problem. Finally, it would be very useful to compare our resulting maps to those

that would be obtained using a simple thresholding method.

Bandeira, L., Marques, J. S., Saraiva, J., and Pina, P. (2011). Automated detection of martian dune fields. *IEEE Geoscience and Remote Sensing Letters*, 8(4):626–630.

Bonnefoy, L. E., Hayes, A. G., Hayne, P. O., Malaska, M. J., Gall, A. L., Solomonidou, A., and Lucas, A. (2016). Compositional and spatial variations in titan dune and interdune regions from cassini {VIMS} and {RADAR}. *Icarus*, 270:222 – 237. Titan's Surface and Atmosphere.

Boser, B. E., Guyon, I. M., and Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. pages 144–152.

Burr, D. M., Drummond, S. A., Cartwright, R., Black, B. A., and Perron, J. T. (2013). Morphology of fluvial networks on titan: Evidence for structural control. *Icarus*, 226(1):742 – 759.

Callegari, M., Casarano, D., Mastrogiuseppe, M., Poggiali, V., and Notarnicola, C. (2014). Dune height estimation on titan exploiting pairs of synthetic aperture radar images with different observation angles. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, PP(99):1–12.

Charnay, B., Barth, E., Rafkin, S., Narteau, C., Lebonnois, S., Rodriguez, S., Courrech du Pont, S., and Lucas, A. (2015). Methane storms as a driver of titan/'s dune orientation. *Nature Geosciences*, 8:362–366.

Ewing, R., Hayes, A., and Lucas, A. (2015). Sand dune patterns on titan controlled by long-term climate cycles. *Nature Geoscience*, 8:15–19.

Le Gall, A., Hayes, A., Ewing, R., Janssen, M., Radebaugh, J., Savage, C., and Encrenaz, P. (2012). Latitudinal and altitudinal controls of titan's dune field morphometry. *Icarus*, 217(1):231 – 242.

Le Gall, A., Janssen, M., Kirk, R., and Lorenz, R. (2014). Modeling microwave backscatter and thermal emission from linear dune fields: Application to titan. *Icarus*, 230(0):198 – 207. Third Planetary Dunes Systems.

Le Gall, A., Janssen, M., Wye, L., Hayes, A., Radebaugh, J., Savage, C., Zebker, H., Lorenz, R., Lunine, J., Kirk, R., Lopes, R., Wall, S., Callahan, P., Stofan, E., and Farr, T. (2011). Cassini sar, radiometry, scatterometry and altimetry observations of titan's dune fields. *Icarus*, 213(2):608 – 624.

Lopes, R. M., Malaska, M., Solomonidou, A., Gall, A. L., Janssen, M., Neish, C., Turtle, E., Birch, S., Hayes, A., Radebaugh, J., Coustenis, A., Schoenfeld, A., Stiles, B., Kirk, R., Mitchell, K., Stofan, E., and Lawrence, K. (2016). Nature, distribution, and origin of titan's undifferentiated plains. *Icarus*, 270:162 – 182. Titan's Surface and Atmosphere.

Lopes, R. M. C., Kirk, R. L., Mitchell, K. L., LeGall, A., Barnes, J. W., Hayes, A., Kargel, J., Wye, L., Radebaugh, J., Stofan, E. R., Janssen, M. A., Neish, C. D., Wall, S. D., Wood, C. A., Lunine, J. I., and Malaska, M. J. (2013). Cryovolcanism on titan: New results from cassini radar and vims. *Journal of Geophysical Research: Planets*, 118(3):416–435.

Lorenz, R. D., Wall, S., Radebaugh, J., Boubin, G., Reffet, E., Janssen, M., Stofan, E., Lopes, R., Kirk, R., Elachi, C., Lunine, J., Mitchell, K., Paganelli, F., Soderblom, L., Wood, C., Wye, L., Zebker, H., Anderson, Y., Ostro, S., Allison, M., Boehmer, R., Callahan, P., Encrenaz, P., Ori, G. G., Francescetti, G., Gim, Y., Hamilton, G., Hensley, S., Johnson, W., Kelleher, K., Muhleman, D., Picardi, G., Posa, F., Roth, L., Seu,

R., Shaffer, S., Stiles, B., Vetrella, S., Flamini, E., and West, R. (2006). The sand seas of titan: Cassini radar observations of longitudinal dunes. *Science*, 312(5774):724–727.

Lucas, A., Aharonson, O., Deledalle, C., Hayes, A. G., Kirk, R., and Howington-Kraus, E. (2014a). Insights into titan's geology and hydrology based on enhanced image processing of cassini radar data. *Journal of Geophysical Research: Planets*, 119(10):2149–2166.

Lucas, A., Rodriguez, S., Narteau, C., Charnay, B., du Pont, S. C., Tokano, T., Garcia, A., Thiriet, M., Hayes, A. G., Lorenz, R. D., and Aharonson, O. (2014b). Growth mechanisms and dune orientation on titan. *Geophysical Research Letters*, 41(17):6093–6100.

McDonald, G. D., Hayes, A. G., Ewing, R. C., Lora, J. M., Newman, C. E., Tokano, T., Lucas, A., Soto, A., and Chen, G. (2016). Variations in titan's dune orientations as a result of orbital forcing. *Icarus*, 270:197 – 210. Titan's Surface and Atmosphere.

Radebaugh, J., Lorenz, R., Farr, T., Paillou, P., Savage, C., and Spencer, C. (2010). Linear dunes on titan and earth: Initial remote sensing comparisons. *Geomorphology*, 121:122 – 132. Planetary Dune Systems.

Radebaugh, J., Lorenz, R., Lunine, J., Wall, S., Boubin, G., Reffet, E., Kirk, R., Lopes, R., Stofan, E., Soderblom, L., Allison, M., Janssen, M., Paillou, P., Callahan, P., Spencer, C., and the Cassini Radar Team (2008). Dunes on titan observed by cassini radar. *Icarus*, 194(2):690 – 703.

Radebaugh, J., Lorenz, R. D., Kirk, R. L., Lunine, J. I., Stofan, E. R., Lopes, R. M., and Wall, S. D. (2007). Mountains on titan observed by cassini radar. *Icarus*, 192(1):77 – 91.

Rodriguez, S., Garcia, A., Lucas, A., Appéré, T., Gall, A. L., Reffet, E., Corre, L. L., Mouélic, S. L., Cornet, T., du Pont, S. C., Narteau, C., Bourgeois, O., Radebaugh, J., Arnold, K., Barnes, J., Stephan, K., Jaumann, R., Sotin, C., Brown, R., Lorenz, R., and Turtle, E. (2014). Global mapping and characterization of titan's dune fields with cassini: Correlation between radar and vims observations. *Icarus*, 230(0):168 – 179. Third Planetary Dunes Systems.

Savage, C. J., Radebaugh, J., Christiansen, E. H., and Lorenz, R. D. (2014). Implications of dune pattern analysis for titan's surface history. *Icarus*, 230(0):180 – 190. Third Planetary Dunes Systems.

Soderblom, L. A., Brown, R. H., Soderblom, J. M., Barnes, J. W., Kirk, R. L., Sotin, C., Jaumann, R., Mackinnon, D. J., Mackowski, D. W., Baines, K. H., Buratti, B. J., Clark, R. N., and Nicholson, P. D. (2009). The geology of hotei regio, titan: Correlation of cassini {VIMS} and {RADAR}. *Icarus*, 204(2):610 – 618.

Soderblom, L. A., Kirk, R. L., Lunine, J. I., Anderson, J. A., Baines, K. H., Barnes, J. W., Barrett, J. M., Brown, R. H., Buratti, B. J., Clark, R. N., Cruikshank, D. P., Elachi, C., Janssen, M. A., Jaumann, R., Karkoschka, E., Mouélic, S. L., Lopes, R. M., Lorenz, R. D., McCord, T. B., Nicholson, P. D., Radebaugh, J., Rizk, B., Sotin, C., Stofan, E. R., Sucharski, T. L., Tomasko, M. G., and Wall, S. D. (2007). Correlations between cassini vims spectra and radar sar images: Implications for titan's surface composition and the character of the huygens probe landing site. *Planetary and Space Science*, 55(13):2025 – 2036. Titan as seen from Huygens.

Stofan, E. R., Elachi, C., Lunine, J. I., Lorenz, R. D., Stiles, B., Mitchell, K. L., Ostro, S., Soderblom, L., Wood, C., Zebker, H., Wall, S., Janssen, M., Kirk, R., Lopes, R., Paganelli, F., Radebaugh, J., Wye, L., Anderson, Y., Allison, M., Boehmer, R., Callahan, P., Encrenaz, P., Flamini, E., Francescetti, G., Gim,

Y., Hamilton, G., Hensley, S., Johnson, W. T. K., Kelleher, K., Muhleman, D., Paillou, P., Picardi, G., Posa, F., Roth, L., Seu, R., Shaffer, S., Vetrella, S., and West, R. (2007). The lakes of titan. *Nature*, 445(7123):61–64.

Turtle, E. P., Perry, J. E., Hayes, A. G., Lorenz, R. D., Barnes, J. W., McEwen, A. S., West, R. A., Del Genio, A. D., Barbara, J. M., Lunine, J. I., Schaller, E. L., Ray, T. L., Lopes, R. M. C., and Stofan, E. R. (2011). Rapid and extensive surface changes near titan's equator: Evidence of april showers. *Science*, 331(6023):1414–1417.