Agricultural Predictions using Satellite Imagery

Sanath N Bhargav
Department of Computer Science
PES University
Bengaluru, Karnataka, India
sanath2705@gmail.com

Sankalp C B
Department of Computer Science
PES University
Bengaluru, Karnataka, India
sankalpbadrinath@gmail.com

Samarth Rajendra
Department of Computer Science
PES University
Bengaluru, Karnataka, India
samarthr007@gmail.com

***Abstract*:**

This paper aims at making agricultural predictions for a specific region in the Indian Subcontinent, by using historical data, thus enabling better agricultural yield. Our study, in this paper is focused on the state of Karnataka. This method uses the data of twelve previous years, taking a monthly average thus countering the seasonality that occurs every ten years. This paper gives predictions using a metric called the normalized difference vegetation index (NDVI). NDVI is the most widely used vegetation index for retrieval of vegetation canopy biophysical properties.

Keywords—component, formatting, style, styling, insert (*key words****)***

I. Introduction

Agriculture, in Karnataka, is one of the oldest occupations, in a way it is embedded deep in the culture of the state. For many people, it serves as their primary means of support. The economy of our state is supported by it. It is the state's main source of income. Exporting crops like cotton, coffee, tea, tobacco, etc., generates foreign exchange. Karnataka is one of the major producers of rice among all other states in India. Hence it is very important to cultivate the right type of crops, in the right place, and at the right time.

Farming as stated before, is one of the main occupations in the state. Many farmers, do not have access to resources that can predict crops they can cultivate, given that they reside in a specific location. This paper provides location-specific predictions, averaged over a square kilometer, facilitating the farmer to cultivate the right type of crop. This paper uses remote sensing data, and extracts the normalized difference vegetation index from spatial satellite imagery. These methods have been implemented before in other regions, but very few have been implemented in Karnataka. This paper aims at providing a facility for the farmers in Karnataka to cultivate their crops in a better way. This in turn reduces the burden on our farmers and enables them to live a better life, by increasing their yield, leading to better profits.

II. Related work

In the state of Karnataka or even in South India, not much research has been conducted on the agricultural predictions using the NDVI index.

Previous researches have used inputs like season type, year of production, area of production, crop type, cloudburst or climatic conditions. They have employed models like the random forest classifier to yield results.

Another paper employs the use of convolutional neural networks to yield results, but it is focused on the regions of the Northern Part of India.

III. Problem Setting and Data Description

Normalized Difference Vegetation Index (NDVI), is a simple graphical indicator that is often used to analyze whether the target being observed contains green health vegetation or not. The NDVI quantifies vegetation by measuring the difference between near-infrared (NIR) and red light-which the vegetation absorbs i.e., has a low reflectance. NDVI values range from +1 to -1, wherein -1 is generally water bodies and +1 is generally dense green-leafy vegetation. Hence, the NDVI is a way of measuring whether a particular location has healthy green vegetation or not. In this paper we have scaled the NDVI value to a range of 10000 to -10000, to enhance readability.

The NDVI values are extracted from satellite imagery, using remote sensing data. The data is accessed manually through the MODIS website of NASA. MODIS stands for Moderate Resolution Imaging Spectroradiometer.

MODIS vegetation indices, produced on a 30-day intervals and at multiple spatial resolutions, provide consistent spatial and temporal comparisons of vegetation canopy greenness, a composite property of leaf-area, chlorophyll and canopy structure. The vegetation indices are retrieved from daily, atmosphere-corrected, bidirectional surface reflectance. The VI's use a MODIS-specific compositing method based on product quality assurance metrics to remove low quality pixels.

From the remaining good quality VI values, a constrained view angle approach then selects a pixel to represent the compositing period (from the two highest NDVI values it selects the pixel that is closest-to-nadir). Because the MODIS sensors aboard Terra and Aqua satellites are identical, the VI algorithm generates each 16-day composite eight days apart (phased products) to permit a higher temporal resolution product by combining both data records. The MODIS VI product suite is now used successfully in all ecosystem, climate, and natural resources management studies and operational research as demonstrated by the ever increasing body of peer publications.

The dataset is prepared by accessing the satellite imagery pertaining to the Indian Subcontinent. The normalized difference vegetation index (NDVI) is calculated using the near infrared (NIR) spectrum and the red spectrum.

The NDVI is calculated using the formula:

$$\text{NDVI} = \frac{(\text{NIR} - \text{Red})}{(\text{NIR} + \text{Red})}$$

The NDVI values are then extracted and stored. The monthly average of the NDVI spanning over a square kilometer is taken.

The aim is to provide NDVI predictions for a given location.

Identify applicable funding agency here. If none, delete this text box.

Fig – Remote Sensing Data from MODIS

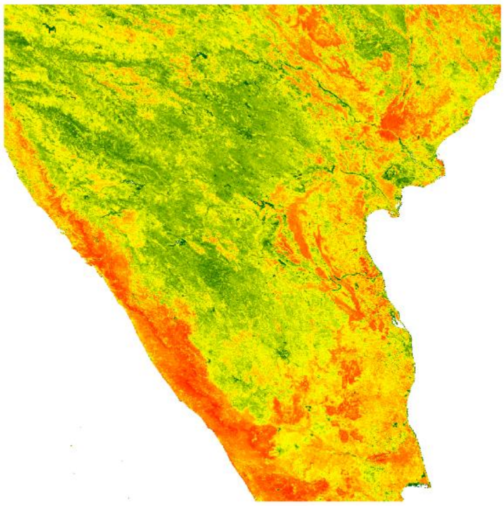

Fig – NDVI

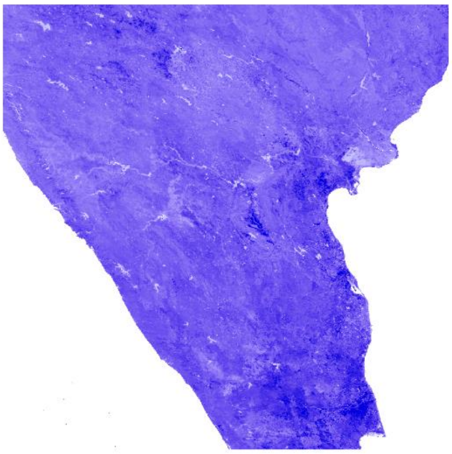

Fig – Near Infrared Band (NIR)

3269	3831	3910	4189	4064	3906	3691	3352	4915	4918	4832	5287	4877	3966	3057	2463	2637	2798	2408	2239
3649	3845	3500	3715	3779	3599	4371	4610	5010	3997	4197	5040	4601	4554	3442	2346	2441	2659	2659	2106
3423	3347	3611	3360	4234	4104	5023	4499	5641	5607	4863	3363	3936	4924	3975	2706	2844	2589	2675	2598
3409	3515	3486	3485	3719	4474	4361	4403	5035	5490	5151	3668	3867	3417	3618	2880	2658	2700	2551	2711
3525	3298	3122	3027	3152	3669	3827	3502	4870	4682	4074	4221	5058	3880	3517	2773	2374	2611	2223	2658
3467	3534	3523	3120	3282	3207	3366	3830	4807	5106	4899	3333	4173	4688	4408	3947	3393	3419	2620	2831
2943	3157	3448	3885	3146	3280	3464	3629	3714	5127	4993	3716	2996	3534	3558	3676	2929	2715	3239	3224
2941	3083	2994	3249	3483	3484	3446	3717	4512	4609	4648	3804	4449	3833	3584	2454	2540	2357	2791	2735
2914	2883	3093	3248	3920	4204	3138	3082	3523	3971	4497	4652	5079	4510	3806	2740	2508	2534	2547	2438
2976	2790	3102	2995	4024	4957	5310	3452	3327	3331	4817	5462	6887	5385	2846	2492	2485	2411	2554	2597
2838	2753	2868	3253	3719	4569	4832	3874	3198	3114	3203	3441	4702	4800	3771	2790	2919	2586	2540	2657
3545	2699	2691	2845	3902	4072	4517	4207	4374	4223	4578	4717	4778	4666	4070	3149	2572	2627	3000	2929
3666	3026	2830	3136	3660	4267	5072	4944	4390	4721	4982	5834	6557	5614	4787	3046	2716	2690	2714	2736
3245	3043	3074	2946	3364	4008	3313	4310	5628	5075	5926	4770	4300	4665	5084	3307	2888	2757	2829	3054
3739	3502	3495	3148	2942	3494	3868	3646	5977	5467	5347	4908	4429	4128	4087	3402	2792	2748	2713	2885
3344	4151	3804	3261	3610	3468	3361	3865	4945	6275	7382	5713	4152	3561	2952	2970	2776	2599	2843	2586
2892	2978	3559	3192	3653	4464	4506	3079	3017	3419	5571	6490	4773	4359	2825	2747	2740	2480	2724	2549
2869	3143	3281	3484	3145	4423	5143	4611	3800	3940	4603	4812	5790	4884	3128	2973	2785	2737	2708	2732
3078	3826	3405	3841	4118	4605	4971	4011	4254	4441	4985	6459	5412	4909	4160	2921	2794	2699	2817	2698
3606	4151	3962	3846	3569	3864	4206	4344	4775	5549	4568	3584	3884	4090	4390	3571	3842	3078	2829	2776
4487	6139	4915	3301	3365	3904	3748	4330	4727	4957	3679	3792	6689	4042	3454	4535	3914	3701	3710	2952
4309	4208	5909	3649	3127	3006	5074	6153	3786	4211	4014	4371	6743	4714	3878	5998	3402	3333	4278	4884
3473	3532	4651	3284	2903	2736	3037	3075	3169	3972	4503	4026	3829	3352	3329	4595	3254	2876	3119	3900
3116	2803	3616	3552	2890	2771	2925	3140	2920	3852	4638	4885	4328	3312	3899	4790	4898	3806	4020	4383
4918	3083	2879	2806	3084	2703	2953	3146	2930	3122	5345	3615	4530	5407	4200	4657	5397	6020	6024	5701
4614	3427	2986	2797	3905	3437	2801	2965	3128	2833	4144	4161	4316	5383	3493	4192	3780	3820	4853	5963
5829	3669	3234	3179	4107	3775	2717	2717	3166	3172	3651	3305	3063	3942	3663	4977	5745	4867	4423	4499

Fig – The extracted NDVI Matrix

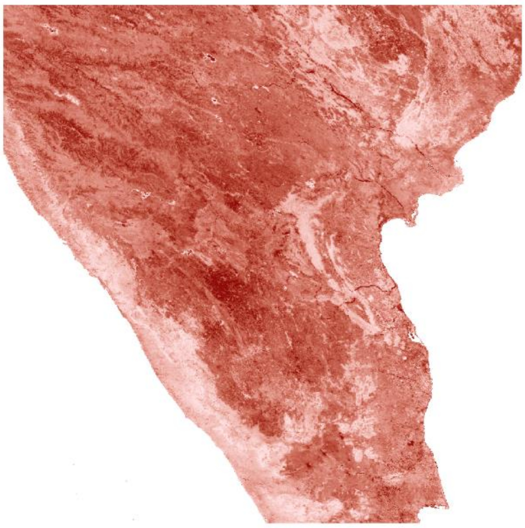

Fig – Red Band

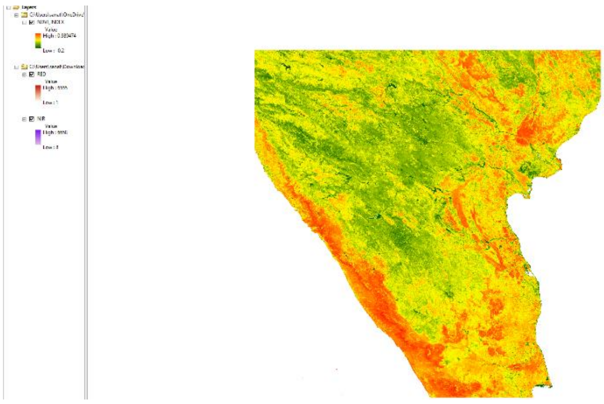

Fig – The NDVI layer formed from the Red and NIR layers

IV. Method Used

The NDVI values gathered from the satellite imagery are stored in an n x n matrix M. For a specific location, the X and Y coordinates of the location are entered and the value of the NDVI corresponding to that location is taken i.e., M[X][Y].

These values are gathered over a period of twelve years, hence 144 values per location in total, stored in a matrix A.

These values are scaled using min-max normalization between 0 and 1 given by:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)}$$

Where x' is the normalized value of A[i]

min(x) is the least value of A

max(x) is the maximum value of A

x is the value A[i]

The normalized values are then stored in a matrix X

A long-short term memory (LSTM) model is used on the training data.

To prepare the data, X is converted into a dataset matrix D, with a time-step of 12 months.

D is then split into D_{train} and D_{test} where the train-test split is 0.8-0.2 respectively.

The LSTM model architecture is as shown below:

The model consists of three sequential LSTM layers, ended by a dense layer to give the output.

Mathematically:

For a single timestep and for a single layer of the LSTM model:

$$f_t = \sigma_g\left(W_f \times x_t + U_f \times h_{t-1} + b_f\right)$$

$$i_t = \sigma_g\left(W_i \times x_t + U_i \times h_{t-1} + b_i\right)$$

$$o_t = \sigma_g\left(W_o \times x_t + U_o \times h_{t-1} + b_o\right)$$

$$c'_t = \sigma_c\left(W_c \times x_t + U_c \times h_{t-1} + b_c\right)$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot c'_t$$

$$h_t = o_t \cdot \sigma_c(c_t)$$

These equations are repeated for twelve time-steps.

The number of layers is 3, hence there are:

6 equations * 3 layers * 12 timesteps = 216 equations.

The

The tanh activation function is applied for each layer of the LSTM where:

$$tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}.$$

The last layer i.e., the dense layer is combined with a mean squared error loss and ADAM optimizer.

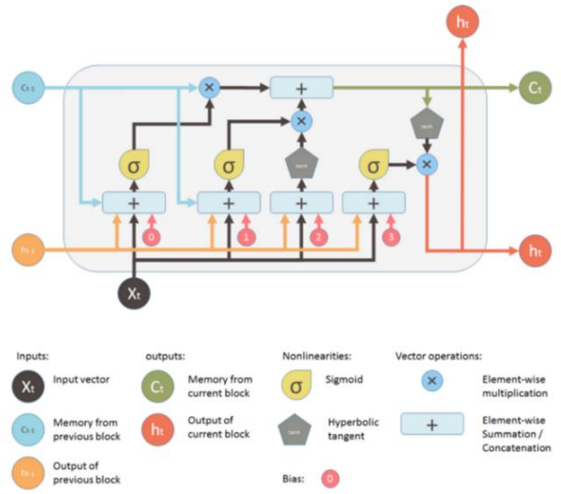

Fig – LSTM Building Block

Model: "sequential_3"

Layer (type)	Output Shape	Param #
lstm_9 (LSTM)	(None, 12, 200)	161600
lstm_10 (LSTM)	(None, 12, 200)	320800
lstm_11 (LSTM)	(None, 200)	320800
dense_3 (Dense)	(None, 1)	201

Total params: 803,401
Trainable params: 803,401
Non-trainable params: 0

Fig – The LSTM model with the corresponding output shapes.

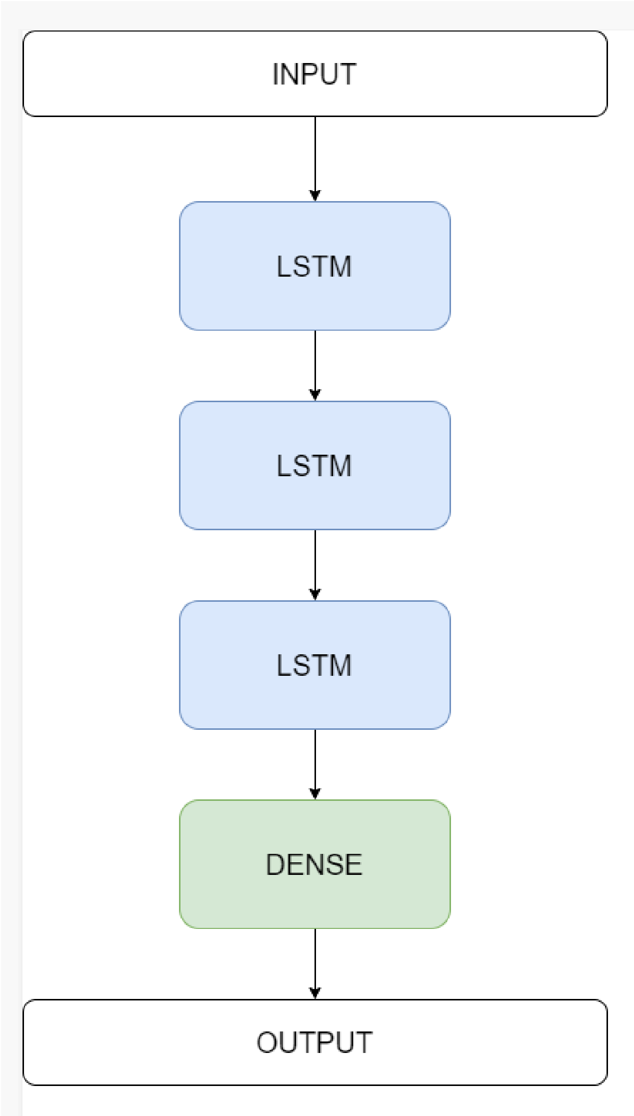

Fig – The LSTM model Architecture

V. Results

The model, when trained over two hundred epochs, gives :

- root mean squared error loss of 0.0844 over the training data
- root mean squared error loss of 0.12 over the test data.

These results are obtained for data collected by monthly averaging the NDVI values over a period of twelve years.

Acknowledgment *(Heading 5)*

We would like to thank PES University for giving us this opportunity to research on this topic. We would like to thank, in particular Prof K S Srinivas, who gave his invaluable guidance, without whom this would be impossible to accomplish.

REFERENCES

1. Ahmadi H, Nusrath A, “Vegetation change Detection of Neka river in Iran by using remote sensing and GIS”, Journal of geography and Geology, 2 (1)., pp. 58-67., 2012
2. Karaburun A. A. K. Bhandari, “Estimation of C factor for soil erosion modelling using NDVI in Buyukcekmece watershed”, Ozean Journal of applied sciences 3, 77-85., 2010.
3. Chouhan R, Rao N," Vegetation detection in Multi spectral remote sensing images: protective Role-analysis of coastal vegetation in 2004 Indian Ocean Tsunami. Geo-Information for Disaster Management ”, Procedia Technology 6, pp., 612 – 621., 2012.
4. Xie Y, Zhao X, Li L, Wang H,” Calculating NDVI for landsat7-etm data after atmospheric Correction using 6s model: a case study in Zhangye City, China, In proceeding IEEE Geo Informatics 18th international conference, pp. 1-4., 2010.
5. Hu Y, Ban Y, Zhang X, Liu J, Zhuang. D,”Spatial-temporal pattern of GIMMS NDVI and its dynamics in Mongolian Plateau”, In IEEE Proceeding on earth observation and remote sensing applications, pp.1-6., 2008.
6. Merwan Achibeta, Stefan Balevb, Antoine Dutotb, Damien Olivierb,” A Model of Road Network and Buildings Extension Co-Evolution Agent-based Modeling and Simulation of Cities Procedia Computer Science”, 32, pp. 828 – 833. 2014
7. Shikhar Deep , Akansha Saklani, ” Urban Sprawl modeling using cellular automata” , The Egyptian Journal of Remote Sensing and Space Sciences, 17, pp. 179–18., 2014.
8. A.K. Bhandari, A. Kumar, “Feature Extraction using Normalized Difference Vegetation Index (NDVI): A Case Study of Jabalpur City”, Proceedings of Communication, Computing & Security. Procedia Technology Volume 6, pp. 612–621, 2012.
9. Nageswara PPR, Shobha SV, Ramesh, KS, Somashekhar RK,” Satellite -based assessment of Agricultural drought in Karnataka State, Journal of the Indian society of remote sensing “, 33 (3), pp. 429-434., 2005
10. Yang Y, Zhu J, Zhao C, Liu S, Tong, The spatial continuity study of NDVI based on Kriging and BPNN algorithm” , Journal of Mathematical and computer modelling., pp. 77 - 85, 2010.
11. Lan Y, Zhang H, Lacey R, Hoffmann WC, Wu W,” Development of an integrated sensor and instrumentation system for measuring crop conditions”, Agricultural engineering journal , 11, pp.11-15., 2009.[12] Bechtel A, Puttmann W, Carlson TN, Ripley DA,” On the relation between NDVI fractional vegetation cover, and leaf area index. Remote Sensing environment, 62 (3), pp. 241 - 252. 1997.
12. El-Shikha DM, Waller P, Hunsaker D, Clarke T, Barnes E,” Ground-based Remote Sensing for assessing water and nitrogen status of broccoli , Agriculture water management, 92, pp. 183-193., 2007.
13. Kim H, Kwak HS, Yoo JS,” Improved clustering algorithm for change detection in remote sensing International journal of digital content technology and its appications, 2008; 2 (2), 55-59.
14. Yamaguchi T, Kishida K, Nunohiro .E, Park JG, Mackin KJ, Matsushita kHK, Harada I., “Artificial neural network paddy-field classifier using Spatiotemporal remote sensing data”, Artificial life and robotics 2010; 15 (2), pp. 221-224
15. .Demirel H, Ozcinar C, Anbarjafari G., “ Satellite image contrast enhancement Using discrete wavelet transform and singular value Decomposition. IEEE Geosciences and remote sensing letters”, 7 (2), pp. 333-337, 2010.
16. Zhang X, Hu Y, Zhuang D, Oi Y, Ma,“NDVI spatial pattern and its differentiation on the Mongolian plateau”, Journal of geographical Sciences, 19, pp., 403- 415, 2009
17. GeoSpatial Analysis 5th Edition, 2015 Smith, Goodchild, Longley. http://www.spatialanalysisonline.com/HTML/?profiles_and_curvature.htm
18. Abbas, G., Ahmad, S., Ahmad, A., Nasim, W., Fatima, Z., Hussain, S., et al. (2017). Quantification the impacts of climate change and crop management on phenology of maize-based cropping system in Punjab, Pakistan. *Agricultural and Forest Meteorology,247,* 42–55.
19. Adger, W. N., Barnett, J., Brown, K., Marshall, N., & O’brien, K. (2013). Cultural dimensions of climate change impacts and adaptation. *Nature Climate Change,3,* 112.
20. Adhikari, S., Baral, H., & Nitschke, C. (2018). Adaptation to climate change in Panchase Mountain ecological regions of Nepal. *Environments,5,* 42.

We suggest that you use a text box to insert a graphic (which is ideally a 300 dpi TIFF or EPS file, with all fonts embedded) because, in an MSW document, this method is somewhat more stable than directly inserting a picture.

To have non-visible rules on your frame, use the MSWord “Format” pull-down menu, select Text Box > Colors and Lines to choose No Fill and No Line.