Land Subsidence Prediction using Recurrent Neural Networks

Sunil Kumar

Indian Institute of Technology (Indian School of Mines): Indian Institute of Technology

Dheeraj Kumar

Indian Institute of Technology (Indian School of Mines): Indian Institute of Technology

Praveen Kumar Donta

Indian Institute of Technology (Indian School of Mines): Indian Institute of Technology

Tarachand Amgoth (✉ tarachand@iitism.ac.in)

Indian Institute of Technology (Indian School of Mines): Indian Institute of Technology

<https://orcid.org/0000-0003-2686-9946>

Research Article

Keywords: Deformation Monitoring, Modified PSInSAR, Recurrent Neural Networks, Vanilla and Stacked LSTM, land Subsidence prediction

Posted Date: July 19th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-278247/v1>

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

[Read Full License](#)

Version of Record: A version of this preprint was published at Stochastic Environmental Research and Risk Assessment on November 24th, 2021. See the published version at <https://doi.org/10.1007/s00477-021-02138-2>.

Land Subsidence Prediction using Recurrent Neural Networks

Sunil Kumar¹ · Dheeraj Kumar¹ · Praveen Kumar Donta² ·
Tarachand Amgoth^{2,*}

Abstract In an environment, one of the natural geological hazards is land surface subsidence. There are several reasons for land subsidence among them are underground coal mining and coal fire in subsurface. The deformation is primarily measured in terms of change in ground elevation values (Z-dimension) at different time intervals at identified ground locations. All the conventional and exiting techniques have certain limitations in monitoring and predicting land surface subsidence. In this work, we predict the land subsidence for one year in the interval of twelve days on the datasets collected through a monitoring technique called Modified PSInSAR. The sample datasets contains 14 locations and 67 previous land subsidence value calculated from each location. We train and test predictive models and perform the prediction of the land subsidence using Vanilla and Stacked long short-term memories (LSTMs). Finally, we demonstrates the predicted deformation values of the 14 locations for one year.

Keywords Deformation Monitoring, Modified PSInSAR, Recurrent Neural Networks, Vanilla and Stacked LSTM · land Subsidence prediction

1 Introduction

In an environment, one of the natural geological hazards is land surface subsidence. There are several reasons for land subsidence among them are underground

coal mining and coal fire in subsurface ([1–4]). Subsidence vulnerability becomes more in those areas where large underground void has been created by extracting coals, ores, etc. ([5–11]). It is very terrible as it involves human losses and great loss of national properties. It also affects the surface and subsurface water resources and the ultimate result is the degradation of the environment. Mine subsidence can take a shape of disaster in inhabited areas if preventive measures are not taken in time. Unfortunately, the increasing demand of energy and mineral resources worldwide has brought mechanization and rapid expansion of mining activities. With such an increase in mining activities, there will be a corresponding increase in mine subsidence problems causing more damage unless proper subsidence control measures are taken. Control measures are directly dependent on detection, monitoring and prediction pattern of subsidence area. Spatial-temporal monitoring and need for precise calculation of land subsidence for mapping in zonal management and corresponding control of surface deformation caused by both underground mining and subsurface fires. However, the effectiveness of the preventative and protective measures of subsidence greatly depends upon the accuracy of subsidence monitoring and associated prediction parameters.

Most of the existing techniques of subsidence monitoring techniques are based on ground-level that rely on the instruments as Precise Level, Auto Level (PL), Digital Level (DL), Total Station (TS), etc. These instruments with associated field survey techniques provide a highly relevant measurement with a millimeter accuracy, but very cumbersome in comparison to modern geo-spatial techniques. Ground-based subsidence monitoring methods are also not safe because measurements are required to be taken along subsidence prone areas. Global Navigation Satellite System (GNSS) systems have made the measurement techniques quite eas-

* Corresponding Author

· ¹Department of Mining Engineering, Indian Institute of Technology (Indian School of Mines), Dhanbad, 826 004, Jharkhand, India.

· ²Department of Computer Science and Engineering, Indian Institute of Technology (Indian School of Mines), Dhanbad, Jharkhand, India.

E-mail: skumar781@iitism.ac.in, dheeraj@iitism.ac.in, praveeniitism@gmail.com, tarachand@iitism.ac.in

Fig. 1: Fringes by DInSAR processed of ALOS POL-SAR Images in mines of JCF

ier in terms of portability of the instruments and satellite dependencies for the data acquisition [12–15]. However, GNSS suffers from some of the same vexing problems as physical movement required in subsidence prone areas with very costly instruments. The limitations of GNSS techniques have been overcome by space-borne imaging techniques. Spaceborne subsidence monitoring has emerged as a better technique after the development of Synthetic Aperture Radar (SAR) Interferometry. The radar satellites can observe almost anywhere on the surface with ease, even during darkness and cloudy conditions, which makes it invaluable for subsidence monitoring. Interferometric Synthetic Aperture Radar (InSAR) technique uses two or more SAR images to generate maps of surface deformation or digital elevation, using differences in the phase of the waves returning to the satellite. Its limitation is accuracy as atmospheric errors introduce several errors. When images are acquired at different times (temporal baseline), utilizing the Differential SAR Interferometry (DInSAR) techniques, it is possible to measure the changes of the surface. These measurements are shown by a series of colored bands, the so-called fringes or interferogram as shown in Fig. 1, are the Angarpathra, Godhar, and Bastacola mines of Jharia coal fields (JCF)¹. The DInSAR technique also has limitations of accuracy and reliability of specific area deformation. Among all the Persistent Scatterer Interferometric Synthetic Aperture Radar (PSInSAR) technique shows better results for deformation of the land surface and with higher accuracy.

The Long short-term memory (LSTMs) are the variants of recurrent neural networks (RNNs) mainly used for time series forecasting. The LSTMs are categorized into Vanilla, Bidirectional, and Stacked LSTMs. The LSTMs have been used to predict the Land subsidence in [16–19] and Bidirectional LSTMs are used in [20, 21]. The LSTMs are efficiently predicting the Land subsi-

dence from the given datasets with better accuracy. It is better to use the Bidirectional LSTMs because it processes the data in both forward and backward time order. The Vanilla and Stacked LSTMs are the variants of LSTMs, where Vanilla LSTMs reduces the memory usage four times better than the LSTMs. So that we can minimize the additional memories during evaluation, and it also speeds up the training and testing. The stacked LSTMs are performing the operations simultaneously using a hierarchy of hidden layers, and it evaluates the complex data very easily with improved accuracy [22, 23].

Our research is based on the processing and analysis of free available SAR datasets as SENTINEL-1A SAR data. These are available on a specific web portal in archive format for the researchers' but lagging by a few months to be available some times. Modified PSInSAR technique is giving large Excel sheets of movement of location termed as PS point and pictorial view with coded colour to identify subsidence on map. Here an excel sheet of 26000 rows and 100 columns had been generated but we ran a sample of 14 rows data sets to study the LSTM module for 1-year prediction. The prediction is for the future so the output of the processed SAR data in Excel format is transferred to the LSTM module for further prediction at 12 days' intervals for a year. This prediction is giving alarming information to deal with settlement and inhabitants residing in the vicinity and saving the life and economy of the nation.

The contributions of this article are summarized as follows:

- Monitor the Land Subsidence data from JCF area using modified PSInSAR and GNSS.
- Train and test the data for accurate predictions of the land surface subsidence using Vanilla and Stacked LSTMs.
- We present the Comparisons of observed and predicted values of land subsidence in JCF of mining as well as GNSS-based locations.
- Finally we also present the predicted values of land surface subsidence for one year.

The remaining sections of this article is arranged as follows. In Section 2, we review some of the existing but related approaches for land subsidence prediction in Mining areas. In Section 3, we provide the preliminaries used in the problem formulation. In section 4, we describe the proposed model in detail. In Section 5, we present the experimental results of the proposed algorithm. This paper is concluded in Section 6.

¹ Jharia coal fields (JCF), Jharkhand state, India

2 Related work

In the recent years Jharia Coal fields (JCF) have witnessed large number of land subsidence and coal fires. Underground activities and coal fires are main cause of such incidents. Nearly 150 such cases have been reported in the past. In [24], Master plan prepared by Bharat Coking Coal Limited (BCCL) in association with central mine planning and design institute (CMPDI) also signifies huge losses in BCCL due to mine fire and induced subsidence in rural as well as urban area of JCF.

The researchers have used conventional DInSAR techniques for monitoring long-term land subsidence phenomenon and achieved the deformation by analyzing the fringes obtained by SAR images ([25–29]). However, the conventional DInSAR techniques have limitations in terms of i) very small spatial baseline ($< 200m$) ii) baseline dependent accuracy of external DEM iii) no reduction of atmospheric phase. To overcome the limitations associated with conventional DInSAR techniques, a first-generation time series InSAR (Advanced DInSAR) technique was introduced by ([30]). Many researchers have worked on advanced DInSAR techniques and developed different approaches to deformation analysis [31–36]. Many studies have been taken out internationally for land subsidence monitoring using Advance DInSAR techniques ([5, 6, 9, 12, 37–40]).

In order to overcome the limitations of second generation advanced DInSAR techniques, the development of the Persistent Scatterer Interferometric Synthetic Aperture Radar (PS-InSAR) technique took place to detect land deformation at the millimeter level. The PSInSAR technique is the geodetic SAR processing technique that uses two or more SAR images to generate maps of topography or deformation of the Earth's surface ([41–43]).

The PSInSAR technique has been applied in the coal filed or nearby areas to detect land deformation using C-Band and L-Band SAR Data ([44, 45]). The short wavelength C-band is more suitable to detect slow velocity subsidence if the optimal baseline is maintained, whereas long-wavelength L-band can effectively detect rapid velocity subsidence ([12, 46, 47]). Its limitation is as less number of PSs achieved due to exclusion of partially correlated scatterers for the analysis resulting in some information losses in the vicinity.

The authors in [48] have developed a modified PS-InSAR approach which is also called Persistent Scatterer Interferometry (PSI). In PSI the correlated scatterers are also included along with permanent scatterers to increase the point target density in the highly susceptible area for decorrelation ([49]). PSI is more reason-

Fig. 2: Shows the GNSS survey points superimposed on LISS IV in Jharia Coal fields, India

able than the Advance DInSAR time arrangement for recognizing most stable dissipating pixels where pixel properties don't fluctuate with time and radar look point ([43, 50].

A study conducted by Central Ground Water Board (CGWB), in Lucknow city, indicated land subsidence is likely to occur due to over exploitation of groundwater in the next 15-20 years if immediate step to increase recharge is not taken, some of the localities of Lucknow in Uttar Pradesh: such as Narhi, Charbagh, Rajajipuram and Gomtinagar regions may see land subsidence by 2026[51].

The authors in [52] have studied the spatial-temporal analysis of land subsidence caused by groundwater pumping from 2010 to 2015 in the Beijing plain using the SBAS InSAR technique. 69 interferograms generated using 47 TerraSAR images were utilized to investigate the land subsidence where long haul groundwater over exploitation and the use of shallow metropolitan space have prompted land subsidence. The greatest yearly land subsidence rate was 146 mm/year from 2011 to 2015. The examination between the SBAS InSAR results and the ground leveling estimations demonstrated that the InSAR land subsidence results accomplished an accuracy of 2 mm. This research work is aimed at the study of the feasibility of the modified PSInSAR technique with C-band SAR data for finding the slow surface deformation caused by coal mine fire and underground mining activities in JCF. Also, a multi-temporal analysis of SAR images of ENVISAT ASAR has been carried out for monitoring and mapping of temporal land subsidence of the area under study. The modified PSI technique has proven its ability to detect land subsidence over the vegetated and rural areas. It also resolves low spatial density of permanent scatterers by considering partially correlated scatterers as permanent

scatterers (PSs) and extracting information from these PSs. The study has been focussed on detecting continuous slow rate subsidence of fifteen major sites of JCF. The imaging techniques also reduce the safety risk and decrease the expenses that are inherent in conventional methods due to extensive fieldwork. In this, the SAR images acquired by SENTINEL-1A of the European Space agency have been processed by SARPROZ Software for deformation analysis of JCF, Dhanbad, India. Further Prediction analysis has been carried out for the determination of vertical shifting of the objects which will be helpful to the safe planning of the projects. In [53, 54], authors have performed theoretical studies how to predict the subsidence above the single and multi-seam longwall mines. However, the work have certain limitations and the prediction is totally based on the physical conditions.

3 Preliminaries

In this section, we provide the preliminaries used in the proposed work. Long Short-Term Memory (LSTM) is used to time series forecasting (TSF). LSTM model that is used for univariate TSF problem is Univariate LSTM. While predict the future values using the past observations through single series of observations and a model is a complicated issue.

As per earlier discussions, the LSTMs operate on sequential data and increasing the number of layers increases the levels of abstraction overtime on the input data. The network depth is more important than the number of memory cells considered for a layer. The LSTM with multiple layers is treated as Stacked LSTM. These upstream layers always provide the sequential output instead of a single output value. Each layer in Stacked LSTMs consider a 3D input for its memory cell and produce a single value as 2D array in an output. Each layer of the stacked LSTM is a chain like processing framework shown in Fig. 3. The cell state in the Fig. 3 is running on top with minor interactions to just

Fig. 3: Structure of a LSTM

flow unchanged data. This layer can remove or add information to the cell state using the two gates such as $\otimes$ and $\oplus$. The σ layer provide the values between 0 and 1 and it is described in Eq. (1).

$$\sigma = \begin{cases} 0 & \text{Nothing through} \\ 1 & \text{Everything through} \end{cases} \quad (1)$$

$$C_{(t-1)} = \begin{cases} 0 & \text{Get rid of complete data} \\ 1 & \text{Keep data completely} \end{cases} \quad (2)$$

The process framework perform majorly four operations to produce the final outcome. First, the forget gate start removing the useless information form the input gate as shown in Eq. 3. It provide input data to the sigmoid function ($\sigma(\dots)$) such as previous layer output i.e. h_{t-1} and hidden layers feature information x_t . The f_t maps the previous layer cell state output i.e. C_{t-1} to the current cell state C_t .

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (3)$$

where W_f is the weight matrix and b_f is the bias vector of the forget gate. The σ is computed as shown in Eq. (4).

$$\sigma(a) = \frac{1}{1 + e^{-a}} \quad (4)$$

In the next state, the LSTM perform two operations $\sigma(\dots)$ and $\tanh(\dots)$ simultaneously to update the current cell state C_t as shown in Eq. (5) and Eq. (6), respectively. The $\sigma(\dots)$ and $\tanh(\dots)$ produces the i_t and $\bar{C}_t$, respectively using previous layer output i.e. h_{t-1} and hidden layers feature information x_t .

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (5)$$

$$\bar{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (6)$$

where W_i and W_c are the weight matrix of input gate and state update vector, respectively, and b_i and b_c are the bias vectors of input gate and state update vector, respectively. The hyperbolic tangent function is represented using $\tanh$ and it is computed as shown in Eq. (7).

$$\tanh(a) = \frac{e^a - e^{-a}}{e^a + e^{-a}} \quad (7)$$

Further, the current state of LSTM updated using previous two operations as shown in Eq. (8).

$$C_t = (C_{t-1} \otimes f_t) + (i_t \otimes \bar{C}_t) \quad (8)$$

where $\otimes$ is point-wise scalar multiplication of the vectors.

Fig. 4: Structure of Stacked LSTM

In the last phase, the LSTM produce the output through output gate by using the current cell state information. The outputs such as O_t and h_t are computed as shown in Eq. (9) and Eq. (10).

$$O_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (9)$$

$$h_t = O_t \otimes \tanh(C_t) \quad (10)$$

In general, the LSTM is comprised of a hidden layer followed by an output layer. Depends on the types of TSF problem, several LSTM techniques are existed in the literature among Vanilla and Stacked LSTMs are more popular. In this work, we tested two univariate LSTM model such as 1) Vanilla LSTM and 2) Stacked LSTM.

3.1 Vanilla LSTM and Stacked LSTM

While making the prediction through Vanilla LSTMs, it uses an output layer and a single hidden LSTM layer. It support the sequence data as an input. The LSTM model reads single time step in each time, unlike the Convolution neural networks (CNN) and the data represented in state format while learning the model.

The LSTM can extend with multiple hidden layer which makes the deeper model called stacked LSTM and which reflects the deep learning model. Increasing the number of layer in the LSTM will increase the prediction accuracy. But, increasing the too many layer also increase the complexity and increases the computational time. The additional layers recombine the representations to provide the new combinations of the predictions with increased abstraction. Instead of giving importance to the memory cells of a layer, considering the depth of network is more important. An LSTM produce the two dimensional output by taking the three dimensional input to the system while performing the learning process. This issue can be addressed by taking the output of each LSTM layer at each time stamp and set the input data with a return *sequences = True*

argument. This allows us to have three dimensional output from hidden LSTM layer as input to the next. Here we build both Vanilla LSTM and Stacked LSTM for predicting the next one Year land subsidence and compared both model with their loss and accuracy. These four steps discussed earlier will repeated over multiple LSTM layers as shown in Fig. 4.

4 Proposed Work

The proposed model is illustrated and depict using flow chart shown in Fig. 5, which is primarily partitioned in to four parts including data collection (along with masking), data augmentation, Training and Evaluation models. Using Vanilla and Stacked LSTMs, predicting land surface subsidence values for one year.

4.1 Data Collection

In this study, we collect the data from 26000 permanent scatterer points generated by SARPROZ software during 67 phases (12 days interval) of SENTINEL-1A images and 58 Global navigation satellite system (GNSS) around the Jharia coal field area in Jharkhand state India. Some important parameters are shown in Table. 1 and locations are pointed on map shown in Fig. 6. The measurement of land surface subsidence for some of the locations in JCF using GNSS are shown in Fig. 7. The JCF is a large coal mines in India with 19.4 billion tonnes of available coal. Since 1916, this area suffered a coal bed fire and consumes nearly 37 millions tons. It results water and air pollution and land subsidence in the city of Jharia [55]. We consider various parameters from the Jharia coalfields including the land type (Agricultural or Barren), transportation facility (Road, Rail, or others), Displacement velocity, depth of the seam and the SAR image availability from 03-10-2016 to 28-12-2018 on every twelve alternative days using the remote sensing. The land subsidence near to the river area is high as compared to the other areas during rainy seasons. In general, the land subsidence patterns have

Fig. 5: Flow Chart of the Proposed Model

Table 1: Dataset locations

Location ID	Latitude	Longitude	Altitude	Location	Displacement velocity (mm/year)	Cumulative displacement (mm)
L1	23.70349	86.42649	182.7	East of barari masjid	-72.6	-162.2
L2	23.70618	86.42627	177.8	Digwadih_NE	-81.8	-182.7
L3	23.70055	86.41453	162.5	Didwadih	-64.7	-144.5
L4	23.73156	86.42495	184.6	Nai dunia basti	-93.8	-209.6
L5	23.78219	86.39222	250.8	Godhar near GDR19	-82	-183.3
L6	23.73856	86.43497	210.3	Laltenbasti	-56.9	-127.1
L7	23.74762	86.32791	203.5	Garbhudi,Moonidih	-27.2	-60.7
L8	23.75829	86.37329	190.4	Balihari,G21	-30	-67
L9	23.75509	86.36283	239.6	Aralgaria,G8	-26.8	-59.9
L10	23.79083	86.35486	210.7	Bansjora,G91	-27.8	-62
L11	23.77601	86.42264	209	Bastacola, nearG12	-30	-67
L12	23.77514	86.38484	220.5	Ghansadih Kendua	-20.6	-46
L13	23.77067	86.36826	216.8	Gopalichak,G5	-25	-55.9
L14	23.8002	86.43213	257.2	Golf ground Dhanbad	2.6	5.8

more diversity among the various locations around the Jharia coalfields, and great challenges to predict the short-term land subsidence. The missing values during

the data collection are masking based on the method used in [56].

Fig. 6: PCs in JCF generated by SARPROZ using processed 67 SAR images of SENTINEL-1A

Fig. 7: The GNSS points of locations L10, L11 and L13

4.2 Data Augmentation and Preparation

For any machine learning methods including stacked LSTMs, training data set is one of the important considerable factors. The best training data will be generated through data augmentation and preparation, and the process is summarized using Fig. 8. The primary goal of the data augmentation is to amplify training data by dividing original data with overlaps to minimize the complexity of the computations. In this work, we consider the time-warping data augmentation approach to prepare the data for training process. We operated with various warping ratios to show the variability of the synthetic training data [57]. With this augmentation process, the training data increases 4-fold with various temporal length.

The data collected and augmented from the remote sensing at GNSS-based locations is one dimensional, where the LSTM requires 3D input data. In data preparation phase, the data initially split in to multiple short sub-sequences and then reshape sub-sequences. We illustrate the data preparation thorough an example for better understand. The learning process of LSTMs follows a sequence of inputs to an output. *An Example of Univariate time series*: {6, 8, 10, 12, 13, 16, 18, 20, 22, 24, 26}. Split these series in to multiple input/output samples. Here, the input can contains more that one step (in this example we consider *three* and represented

using X) and an output as one step (treated as Y) as shown below:

X	Y
[6 8 10]	12
[8 10 12]	13
[10 12 13]	16
...	...
[20 22 24]	26

In this way, the input data split and reshare the data for providing the input to training or testing modules. The input data categorized into training and test data. The training data used in Stacked LSTM learning models where as the test data used for predicting the results.

4.3 Training and Evaluation Model

The training data augmented and prepared for input to the Stacked LSTM. The over-fitting is avoided by adopting the dropout strategy after each LSTM for better generalization. The dropout rate is approximately 10% [58]. The stacked LSTM have the capability to handle the long and short-term time dependencies for forecasting the accurate land subsidence. The first LSTM layer send a sequence vector to the next LSTM layer and so on. Each subsequent LSTM receives previous time stamp's feedback that either allow for process or drop the data. The basic working model of the stacked LSTM is presented in Section 3.

The test data prepared for testing and no augmentation performed on it. After the model being learned, a test data set is taken as an input to evaluate the model efficiency to validate the accuracy of the predicted values.

5 Experimental Results

In this section we evaluate the prediction accuracy of the land subsidence through Vanilla LSTM and Stacked LSTMs. Initially, we discuss the model construction followed by the results analysis.

5.1 Model Construction

In this work, we collect the data from 14 locations around Jharia coal field from Jharkhand state in India between 3rd Oct 2016 to 28th Dec 2018, i.e. 817 days. We consider 14 environmental conditions and each condition 817 samples are collected. All these conditions are segmented with the length of 128 and 60% overlap. In this work, we use 64, 32, and 32 hidden units

Fig. 8: Process Model of Data Augmentation and Preparation

in LSTM layer 1, layer 2 and layer 3, respectively. The input layer has 128 units which is equal to input sample dimension. We can fit the training dataset once the complete model is defined. For the whole dataset, 70% is used for training and remaining used for testing. The Root Mean Square Error (RMSE) is an indicator to evaluate the performance of training model.

5.2 Results Analysis

In this section, we compare the simulation results of Stacked LSTM and Vanilla LSTM using various metrics such as accuracy, land subsidence, and etc. The proposed approach is used regression model, so Root mean squared error is a good measure of accuracy.

5.2.1 Root Mean Square Errors

It is measured as the difference between values predicted by a model and the values observed and it is denoted using $\mathcal{R}_m$. It is calculated using Eq. (11)

$$\mathcal{R}_m = \sqrt{\frac{\sum_{i=1}^N (X_i - \hat{X}_i)^2}{N}} \quad (11)$$

where X_i is the actual data and $\hat{X}_i$ is the predicted data of data set i , and N indicates the number of samples.

5.2.2 Root Mean Absolute Percentage Errors

It is measured as the difference between values predicted by a model and the values observed and it is denoted as $\mathcal{R}_p$. It is calculated using Eq. (11)

$$\mathcal{R}_p = \frac{100}{\sqrt{N}} \times \sqrt{\sum_{i=1}^N \left(\frac{X_i - \hat{X}_i}{X_i} \right)^2} \quad (12)$$

Table 2: Comparisons of Vanilla LSTM and Stacked LSTM Concerning $\mathcal{R}_m$ and $\mathcal{R}_p$

Location	Vanilla LSTM		Stacked LSTM	
	$\mathcal{R}_m$	$\mathcal{R}_p$	$\mathcal{R}_m$	$\mathcal{R}_p$
L1	0.0817	71.4	0.0139	88.204
L2	0.062	74.96	0.038	80.364
L3	0.059	75.68	0.050	77.639
L4	0.053	76.85	0.0315	82.227
L5	0.058	75.91	0.049	77.655
L6	0.069	73.62	0.037	80.5
L7	0.147	61.604	0.078	71.9
L8	0.106	67.31	0.15	61.03
L9	0.102	68.00	0.08	71.4
L10	0.091	69.7	0.052	77.09
L11	0.117	65.697	0.0512	77.35
L12	0.067	74.01	0.047	78.28
L13	0.055	76.46	0.0401	79.97
L14	0.163	59.51	0.09	68.4

where X_i is the actual data and $\hat{X}_i$ is the predicted data of data set i , and N indicates the number of samples. The comparison results of running example dataset in terms of $\mathcal{R}_m$ and $\mathcal{R}_p$ are presented in Table. 2

The comparison between the observed and predicted train and test land subsistence (reduced level) values of Location L1 to L14 with Vanilla LSTM and Stacked LSTM model are presented in Fig. 9 to Fig. 11. We observe that the prediction accuracy of both the models is between 80% to 95% for 14 locations. We estimate the deformation of next one year with every twelve alternative days and the predicted deformation values (mm) are presented in Table. 3. As the analysis show the subsidence in Nai dunia basti near Jharia (L4) is alarming as 93.8 mm /year where as Digwadih(L2) and Godhar(L5) is also showing critical rate as 82 mm/year during the period 2016-19. Based on the prediction result Godhar (L5) is showing alarming as 105 mm/year and L4 & L2 is showing 97 mm/year and 71 mm/year respectively.

Fig. 9: Comparison between the observed and predicted values of land subsidence of Location L1 to L6 with Vanilla and Stacked LSTM model

Fig. 10: Comparison between the observed and predicted values of land subsidence of each Location L7 to L12 with Vanilla and Stacked LSTMs model

Fig. 11: Comparison between the observed and predicted values of land subsidence of each Locations L13 and L14 with Vanilla and Stacked LSTMs model

5.3 Hyperparameters' Influence

In this work, two hyperparameters are highly influenced on the accuracy and Time consumption which are (a) #of LSTM layers and (b) Size of the input. We conduct experiments on these hyperparameters and plot the results in Fig. 12 and Fig. 13.

In Fig. 12, we evaluate the accuracy by considering the two hyperparameters. In Fig. 12(a), we assess the accuracy of fourteen data labels by increasing the # of LSTM layers in the stacked LSTM. We notice that increasing the number of layers also increases the accuracy. Similarly, we evaluate the accuracy by increasing the input data's size, and the results are plot in Fig. 12(b). Here, we observe that increasing the input size also increases the accuracy. In Fig. 13, we evaluate the computation time per epoch by considering the two hyperparameters. In Fig. 13(a), we assess the time computation of fourteen data labels by increasing the # of LSTM layers in the stacked LSTM. We notice that increasing the number of layers also increases the computation time slightly. We conclude that increasing the layers in the stacked LSTM also increases the computational time. Similarly, we evaluate the time computation by increasing the input data's size, and the results are plot in Fig. 13(b). Here, we observe that increasing the input size also increases the computational time. In the proposed work, we used three layers to balance the prediction accuracy and computational time.

6 Conclusion

In this work, we have presented a new scientific approach for monitoring and prediction of mining induced land subsidence in Jharia Coalfield using modified PSInSAR, GNSS and Recurrent Neural Networks. We have used two variants of RNN are Vanilla LSTM and Stacked LSTM. We have collected 67 datasets pertaining to land subsidence at 14 various locations in JCF at an interval of 12 days. We have used modified PSInSAR technique to collect the land subsidence value. To perform prediction of land subsidence, we train, test, and validate the predictive models by splitting datasets into 7:2:1. Finally, we have predicted the land subsidence for one year i.e, next 30 predictions in the interval of 12 days and demonstrated the prediction deformation values of all the 14 locations.

Acknowledgement

The authors would like to extends warm thanks to Dr. Daniele Perissin for providing access to SARPROZ software for this research study. The authors convey sincere thanks to the officials of the Department of Environment and Department of Surveying, BCCL, Dhanbad, India, for providing supports during the field studies.

References

1. Sunil Kumar, Dheeraj Kumar, Sumit Kumar Chaudhary, Narendra Singh, and Kapil Kumar Ma-

Table 3: Predicted Deformation values (mm) using Stacked LSTM

Day (X 12)	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12	L13	L14
68	-165	-181	-150.1	-224	-202.4	-111	-59	-62.5	-66	-78	-68	-49	-48	-8
69	-172	-183.1	-164	-225	-212	-116	-60	-66	-70	-79	-70	-50	-54.2	-8.3
70	-175	-187	-157.2	-232	-218	-118	-59.6	-65	-83	-82	-70.2	-48	-52	-10
71	-176	-190	-161.3	-233	-224.1	-119	-59.5	-66.2	-82.2	-85.2	-75	-51	-51	-11
72	-178	-193.2	-166.1	-231	-227.2	-121	-61.4	-64	-78	-83	-72.5	-55	-53	-11.2
73	-201	-195	-157.8	-240.1	-229	-123	-62	-65	-84	-86	-74	-54	-58	-11.4
74	-202	-199	-160.3	-246.3	-238.2	-126	-61.8	-62	-81	-90.1	-73	-54.7	-55	-11
75	-204.5	-200.2	-167.2	-249	-236	-129	-63	-67	-83.8	-89	-75.1	-56	-56	-10.6
76	-205.7	-203	-166	-251	-240	-131.5	-64	-65	-84	-93	-76	-57.8	-58.1	-12
77	-207	-205	-180	-256	-238	-137.1	-67	-67	-85	-93.2	-76.6	-57	-59	-12.2
78	-209	-208	-181.1	-262	-241	-135.3	-65	-70	-84.6	-94	-77	-56	-60.4	-14
79	-211.4	-211.1	-175	-271	-245.2	-140.1	-61.8	-72	-90	-96	-78	-62	-60.6	-11
80	-214	-215	-185.2	-264	-251.1	-142.6	-69	-73	-91	-99.1	-80	-61	-62	-12.6
81	-214.5	-217	-181.1	-267	-263	-141	-75	-74	-101.7	-97	-81	-60	-63	-11
82	-212.5	-216	-177	-274.2	-255	-147.1	-78	-81	-102	-97.5	-82	-62.2	-63.2	-14
83	-217.3	-218	-176	-278	-265.5	-146	-78.2	-83	-108	-95	-84	-63	-66	-14.2
84	-218.1	-220	-189	-282	-259	-150.5	-78.1	-85	-105	-98	-82	-64	-66.5	-12
85	-221	-222.1	-191	-286	-264	-153.8	-78.5	-86	-115	-101	-88	-61	-62	-17
86	-223	-225	-188	-285	-268	-152	-82	-89	-130	-99.2	-90	-68	-66	-18
87	-224	-227.6	-199.2	-288	-270	-151.2	-81	-102	-129	-102	-91	-67	-68	-16
88	-223	-229	-201	-292	-273	-156	-80	-85	-127	-98.1	-97	-68	-68.5	-12
89	-226	-231	-203	-291	-278	-161	-84	-101	-131	-101.6	-100	-69	-68	-13
90	-229	-234	-207	-296	-289	-177	-83	-81	-129.2	-109	-109	-74	-72	-12.5
91	-232	-237	-205	-298	-285	-172	-78	-89	-130	-112	-112	-68	-73.2	-13.8
92	-233.6	-239	-213	-301	-290	-164	-78.3	-93	-127	-116	-113	-71	-74	-14.5
93	-225.4	-241	-212.1	-308	-295	-166	-74	-100	-128	-117	-112	-71	-76	-14.6
94	-227.2	-245	-209.8	-307	-296	-168.5	-73.1	-105	-131	-115	-115	-65	-76.4	-16.6
95	-229	-245.3	-213	-315	-301.5	-169	-78	-109	-134.2	-114	-117	-72.2	-77	-16.9
96	-235	-248	-215	-317	-305	-172	-77	-101	-135	-118	-119	-74	-78	-17
97	-236	-251	-216	-318	-306.1	-173.1	-79	-106	-134.6	-120	-120	-76	-76	-19
98	-234	-252	-217	-321	-307	-174	-80.1	-114	-137	-120.1	-122	-79	-80	18.6

lik. Land subsidence mapping and monitoring using modified persistent scatterer interferometric synthetic aperture radar in jharia coalfield, india. *Journal of Earth System Science*, 129(1):1–10, 2020.

2. Jitendra Pandey, Dheeraj Kumar, Virendra Kumar Singh, and Nitroj Kumar Mohalik. Environmental and socio-economic impacts of fire in jharia coalfield, jharkhand, india: an appraisal. *Curr. Sci*, 110(9):1639–1650, 2016.

Fig. 12: Influence of Accuracy with varying hyperparameters (a) #of LSTM Layers (b) Size of input

Fig. 13: Influence of Time computation with varying hyperparameters (a) #of LSTM Layers (b) Size of input

3. RS Chatterjee. Coal fire mapping from satellite thermal ir data—a case example in jharia coalfield, jharkhand, india. *ISPRS Journal of Photogrammetry and Remote Sensing*, 60(2):113–128, 2006.
4. Song Jianjun, Han Chunjian, Li Ping, Zhang Junwei, Liu Deyuan, Jiang Minde, Zheng Lin, Zhang Jingkai, and Song Jianying. Quantitative prediction of mining subsidence and its impact on the environment. *International Journal of Mining Science and Technology*, 22(1):69–73, 2012.
5. SG Ishwar and Dheeraj Kumar. Application of din-sar in mine surface subsidence monitoring and prediction. *Current Science*, 112(1):46–51, 2017.
6. Yuxiao Qin and Daniele Perissin. Monitoring underground mining subsidence in south india with c-and l-band insar technique. In *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 294–297. IEEE, 2015.
7. Jeanine Engelbrecht and Michael Inngs. Differential interferometry techniques on l-band data employed for the monitoring of surface subsidence due

- to mining. *South African Journal of Geomatics*, 2 (2):82–93, 2013.
8. Waldir R Paradella, Alessandro Ferretti, José C Mura, Davide Colombo, Fabio F Gama, Andrea Tamburini, Athos R Santos, Fabrizio Novali, Mauricio Galo, Paulo O Camargo, et al. Mapping surface deformation in open pit iron mines of carajás province (amazon region) using an integrated sar analysis. *Engineering Geology*, 193: 61–78, 2015.
 9. Mukesh Gupta, Kamini Kanta Mohanty, Devendra Kumar, and Ramanuj Banerjee. Monitoring surface elevation changes in jharia coalfield, india using synthetic aperture radar interferometry. *Environmental earth sciences*, 71(6):2875–2883, 2014.
 10. Liu Guang, Guo Huadong, Guo Xiaofang, Zbigniew Perski, Yue Huanyin, et al. Mining area subsidence monitoring using multi-band sar data. In *2009 Joint Urban Remote Sensing Event*, pages 1–6. Ieee, 2009.
 11. Fang Miao, Mingxing Yan, Xiaoying Qi, Chengming Ye, Baocun Wang, Rui Liu, and Jianhua Chen. Application of dinsar and gis for underground mine subsidence monitoring. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, 37:251–255, 2008.
 12. RS Chatterjee, Shailaja Thapa, KB Singh, G Varunakumar, and EVR Raju. Detecting, mapping and monitoring of land subsidence in jharia coalfield, jharkhand, india by spaceborne differential interferometric sar, gps and precision levelling techniques. *Journal of Earth System Science*, 124 (6):1359–1376, 2015.
 13. Jian Wang, Xiangguo Peng, and Chang hui Xu. Coal mining gps subsidence monitoring technology and its application. *Mining Science and Technology (China)*, 21(4):463–467, 2011.
 14. Gao Jing-Xiang and Hu Hong. Advanced gnss technology of mining deformation monitoring. *Procedia Earth and Planetary Science*, 1(1):1081–1088, 2009.
 15. Wei-Cai Lü, Shi-Guang Cheng, Hai-Sheng Yang, and Da-Peng Liu. Application of gps technology to build a mine-subsidence observation station. *Journal of China University of Mining and Technology*, 18(3):377–380, 2008.
 16. Huijun Li, Lin Zhu, Huili Gong, Hanrui Sun, and Jie Yu. Land subsidence modelling using a long short-term memory algorithm based on time-series datasets. *Proceedings of the International Association of Hydrological Sciences*, 382:505–510, 2020.
 17. Yi-Ling Qiao, Yu-Kun Lai, Hongbo Fu, and Lin Gao. Synthesizing mesh deformation sequences with bidirectional LSTM. *IEEE Transactions on Visualization and Computer Graphics*, 2020.
 18. Mehreen Mubashar, Gul Muhammad Khan, and Ramla Khan. Landslide prediction using long short term memory (LSTM) neural network on time series data in pakistan. In *2021 International Conference on Artificial Intelligence (ICAI)*, pages 175–181. IEEE, 2021.
 19. Fangling Pu, Zhaozhuo Xu, Hongyu Chen, Xin Xu, and Nengcheng Chen. A dlm-lstm framework for north-south land deformation trend analysis from low-cost gps sensor time series. *Journal of Sensors*, 2018, 2018.
 20. Shui-Long Shen, Pierre Guy Atangana Njock, Annan Zhou, and Hai-Min Lyu. Dynamic prediction of jet grouted column diameter in soft soil using Bi-LSTM deep learning. *Acta Geotechnica*, 16(1): 303–315, 2021.
 21. Xudong Qu, Jie Yang, and Meng Chang. A deep learning model for concrete dam deformation prediction based on RS-LSTM. *Journal of Sensors*, 2019, 2019.
 22. Tryambak Gangopadhyay, Sin Yong Tan, Genyi Huang, and Soumik Sarkar. Temporal attention and stacked lstms for multivariate time series prediction. *NIPS 2018 Workshop Spatiotemporal Blind Submission*, 2018.
 23. Sisheng Liang, Long Nguyen, and Fang Jin. A multi-variable stacked long-short term memory network for wind speed forecasting. In *2018 IEEE International Conference on Big Data (Big Data)*, pages 4561–4564. IEEE, 2018.
 24. BCCL. Master plan for dealing with fire, subsidence and rehabilitation in the leasehold of bccl, 2008.
 25. RS Chatterjee, Benedicte Fruneau, JP Rudant, PS Roy, Pierre-Louis Frison, RC Lakhera, VK Dadhwal, and Ranajit Saha. Subsidence of kolkata (calcutta) city, india during the 1990s as observed from space by differential synthetic aperture radar interferometry (d-insar) technique. *Remote Sensing of Environment*, 102(1-2):176–185, 2006.
 26. Tazlo Strozzi, Urs Wegmuller, Luigi Tosi, Gabrlele Bitelli, and Volker Spreckels. Land subsidence monitoring with differential sar interferometry. *Photogrammetric engineering and remote sensing*, 67 (11):1261–1270, 2001.
 27. Riccardo Lanari, Paul Lundgren, Mariarosaria Manzo, and Francesco Casu. Satellite radar interferometry time series analysis of surface deformation for los angeles, california. *Geophysical Research Letters*, 31(23), 2004.
 28. Daniel Raucoules, Christophe Maisons, Claudie Carnec, Stéphane Le Mouelic, Christine King, and

- Steven Hosford. Monitoring of slow ground deformation by ers radar interferometry on the vauvert salt mine (france): Comparison with ground-based measurement. *Remote sensing of environment*, 88 (4):468–478, 2003.
29. Falk Amelung, Sigurjón Jónsson, Howard Zebker, and Paul Segall. Widespread uplift and ‘trap-door’ faulting on galapagos volcanoes observed with radar interferometry. *Nature*, 407(6807):993–996, 2000.
 30. Alessandro Ferretti, Claudio Prati, and Fabio Rocca. Permanent scatterers in sar interferometry. *IEEE Transactions on geoscience and remote sensing*, 39(1):8–20, 2001.
 31. Andrew Hooper, Howard Zebker, Paul Segall, and Bert Kampes. A new method for measuring deformation on volcanoes and other natural terrains using insar persistent scatterers. *Geophysical research letters*, 31(23), 2004.
 32. Andrew Hooper and Howard A Zebker. Phase unwrapping in three dimensions with application to insar time series. *JOSA A*, 24(9):2737–2747, 2007.
 33. Andrew Hooper. A multi-temporal insar method incorporating both persistent scatterer and small baseline approaches. *Geophysical Research Letters*, 35(16), 2008.
 34. M Crosetto, B Crippa, E Biescas, O Monserrat, M Agudo, and P Fernández. Land deformation measurement using sar interferometry: state-of-the-art. *Photogrammetrie Fernerkundung Geoinformation*, 2005(6):497, 2005.
 35. Oscar Mora, Jordi J Mallorqui, and Antoni Broquetas. Linear and nonlinear terrain deformation maps from a reduced set of interferometric sar images. *IEEE Transactions on Geoscience and Remote Sensing*, 41(10):2243–2253, 2003.
 36. David A Schmidt and Roland Bürgmann. Time-dependent land uplift and subsidence in the santa clara valley, california, from a large interferometric synthetic aperture radar data set. *Journal of Geophysical Research: Solid Earth*, 108(B9), 2003.
 37. Shaochun Dong, Hongwei Yin, Suping Yao, and Fei Zhang. Detecting surface subsidence in coal mining area based on dinsar technique. *Journal of Earth Science*, 24(3):449–456, 2013.
 38. RS Chatterjee, KB Singh, Shailaja Thapa, and Dheeraj Kumar. The present status of subsiding land vulnerable to roof collapse in the jharia coal-field, india, as obtained from shorter temporal baseline c-band dinsar by smaller spatial subset unwrapped phase profiling. *International Journal of Remote Sensing*, 37(1):176–190, 2016.
 39. Jeanine Engelbrecht, Michael Raymond Inggs, and G Makusha. Detection and monitoring of surface subsidence associated with mining activities in the witbank coalfields, south africa, using differential radar interferometry. *South African Journal of Geology*, 114(1):77–94, 2011.
 40. Maria Przyłucka, Gerardo Herrera, Marek Graniczny, Davide Colombo, and Marta Béjar-Pizarro. Combination of conventional and advanced dinsar to monitor very fast mining subsidence with terrasar-x data: Bytom city (poland). *Remote Sensing*, 7(5):5300–5328, 2015.
 41. CLAUDIO Prati, Alessandro Ferretti, and Daniele Perissin. Recent advances on surface ground deformation measurement by means of repeated spaceborne sar observations. *Journal of Geodynamics*, 49(3-4):161–170, 2010.
 42. Andrew Hooper, David Bekaert, Karsten Spaans, and Mahmut Arıkan. Recent advances in sar interferometry time series analysis for measuring crustal deformation. *Tectonophysics*, 514:1–13, 2012.
 43. José Claudio Mura, Waldir R Paradella, Fabio F Gama, Guilherme G Silva, Mauricio Galo, Paulo O Camargo, Arnaldo Q Silva, and Aristotelina Silva. Monitoring of non-linear ground movement in an open pit iron mine based on an integration of advanced dinsar techniques using terrasar-x data. *Remote Sensing*, 8(5):409, 2016.
 44. Saygın Abdikan, Mahmut Arıkan, Fusun Balik Sanlı, and Ziyadin Cakir. Monitoring of coal mining subsidence in peri-urban area of zonguldak city (nw turkey) with persistent scatterer interferometry using alos-palsar. *Environmental Earth Sciences*, 71 (9):4081–4089, 2014.
 45. Shailaja Thapa, RS Chatterjee, KB Singh, and Dheeraj Kumar. Land subsidence monitoring using ps-insar technique for l-band sar data. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 40, 2016.
 46. Saygın Abdikan, Andrew Hooper, Mahmut Arıkan, F Balik Sanlı, Ziyadin Cakir, and Hüseyin Kemaldere. Insar time series analysis of coal mining in zonguldak city, northwestern turkey, 2011.
 47. Huanyin Yue, Guang Liu, Zbigniew Perski, and Huadong Guo. Satellite radar reveals land subsidence over coal mines. *SPIE newsroom*, 3, 2011.
 48. Daniele Perissin and Teng Wang. Repeat-pass sar interferometry with partially coherent targets. *IEEE Transactions on Geoscience and Remote Sensing*, 50(1):271–280, 2011.
 49. Umut G Sefercik and Uwe Soergel. Influence of temporal baseline on the vertical absolute accuracies of tsx hs interferometric dsms: Case study on

- berlin. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 2014.
50. Michele Crosetto, Oriol Monserrat, María Cuevas-González, Núria Devanthéry, and Bruno Crippa. Persistent scatterer interferometry: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 115:78–89, 2016.
 51. Trivedi. Land subsidence threat by 2026, 2020.
 52. Chaofan Zhou, Huili Gong, Beibei Chen, Jiwei Li, Mingliang Gao, Feng Zhu, Wenfeng Chen, and Yue Liang. Insar time-series analysis of land subsidence under different land use types in the eastern beijing plain, china. *Remote Sensing*, 9(4):380, 2017.
 53. AM Suchowerska Iwanec, JP Carter, and JP Hambleton. Geomechanics of subsidence above single and multi-seam coal mining. *Journal of Rock Mechanics and Geotechnical Engineering*, 8(3):304–313, 2016.
 54. The SAR PROcessing tool by periZ. <http://www.sarproz.com>, 2019. Accessed: June 30, 2021.
 55. Sandeep Pai and Savannah Carr-Wilson. *Total Transition: The Human Side of the Renewable Energy Revolution*. Rocky Mountain Books Ltd, 2018.
 56. Zhengping Che, Sanjay Purushotham, Kyunghyun Cho, David Sontag, and Yan Liu. Recurrent neural networks for multivariate time series with missing values. *Scientific reports*, 8(1):1–12, 2018.
 57. Khandakar M Rashid and Joseph Louis. Times-series data augmentation and deep learning for construction equipment activity recognition. *Advanced Engineering Informatics*, 42:100944, 2019.
 58. Chih-Chiang Wei. Development of stacked long short-term memory neural networks with numerical solutions for wind velocity predictions. *Advances in Meteorology*, 2020, 2020.