

# Accepted Article

## Time-Series Prediction Approaches to Forecasting Deformation in Sentinel-1 InSAR Data

P. Hill<sup>1</sup>, J. Biggs<sup>2</sup>, V. Ponce-López<sup>1</sup> and D. Bull<sup>1</sup>,

<sup>1</sup>Department of Electrical and Electronic Engineering, University of Bristol, Bristol, United Kingdom

<sup>2</sup>COMET, School of Earth Sciences, University of Bristol, Bristol, United Kingdom

### Key Points:

- We test established time series prediction methods on 4 years of Sentinel-1 InSAR data, and investigate the role of seasonality.
- For seasonal signals, SARIMA and machine learning (LSTM) perform best over <3 months, and sinusoid extrapolation over >6 months.
- Forecast quality decreases for less seasonal signals, and a constant value prediction performs best for randomly-selected datapoints.

---

Corresponding author: Paul Hill, [Paul.Hill@Bristol.ac.uk](mailto:Paul.Hill@Bristol.ac.uk)

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1029/2020JB020176](https://doi.org/10.1029/2020JB020176).

This article is protected by copyright. All rights reserved.

13      **Abstract**

14      Time series of displacement are now routinely available from satellite InSAR and are used  
 15      for flagging anomalous ground motion, but not yet forecasting. We test conventional time  
 16      series forecasting methods such as SARIMA and supervised machine learning approaches  
 17      such as LSTM compared to simple function extrapolation. We focus initially on fore-  
 18      casting seasonal signals and begin by characterising the time-series using sinusoid fitting,  
 19      seasonal decomposition and autocorrelation functions. We find that the three measures  
 20      are broadly comparable but identify different types of seasonal characteristic. We use  
 21      this to select a set of 310 points with highly seasonal characteristics and test the three  
 22      chosen forecasting methods over prediction windows of 1-9 months. The lowest overall  
 23      median RMSE values are obtained for SARIMA when considering short term predictions  
 24      (<1 month), whereas sinusoid extrapolation produces the lowest median RMSE values  
 25      for longer predictions (>6 months). Machine learning methods (LSTM) perform less well.  
 26      We then test the prediction methods on 2000 randomly selected points with a range of  
 27      seasonalities and find that simple extrapolation of a constant function performed bet-  
 28      ter overall than any of the more sophisticated time series prediction methods. Compar-  
 29      isons between seasonality and RMSE show a small improvement in performance with in-  
 30      creasing seasonality. This proof-of-concept study demonstrates the potential of time-series  
 31      prediction for InSAR data but also highlights the limitations of applying these techniques  
 32      to non-periodic signals or individual measurement points. We anticipate future devel-  
 33      opments, especially to shorter timescales, will have a broad range of potential applica-  
 34      tions, from infrastructure stability to volcanic eruptions.

35      **1 Introduction**

36      Many tectonically stable regions suffer from significant ground motion due to the effects  
 37      of former coalfields (McCay et al., 2018), landslides (Chambers et al., 2008), the shrink  
 38      and swell of shallow clays (Crilly, 2001; Aldiss et al., 2014), tree growth, coastal erosion,  
 39      natural sinkholes (Lamont-Black et al., 2002; Banks et al., 1995) and tunnelling (e.g. Cross-  
 40      rail, (Milillo et al., 2018)). Ground motion analysis has recently focused on satellite-based  
 41      InSAR, which uses the phase difference between pairs of radar satellite images to map  
 42      ground deformation at mm/yr precision. In particular, the Copernicus Sentinel-1 con-  
 43      stellation has revolutionised the coverage, frequency and availability of InSAR data and  
 44      can be used to produce high-resolution maps of ground motion across Europe every six  
 45      days in near real-time. To this end, many companies have generated post-processed ground  
 46      motion data maps and time series based on Sentinel-1 InSAR data (e.g. [cggs.com](http://cggs.com); [sat-sense.com](http://sat-sense.com); [tre-altamira.com](http://tre-altamira.com)). Machine learning methods have been used to automati-  
 47      cally flag deformation, or changes in deformation in the large datasets (Anantrasirichai  
 48      et al., 2018, 2019a, 2019b; Gaddes et al., 2018, 2019; Valade et al., 2019). Here we in-  
 49      vestigate the possibility that these Sentinel-1 datasets can be used to forecast future be-  
 50      haviour.

51      Time series forecasting defines a prediction model to forecast future values of a uni-  
 52      variate or multivariate time series based on previously observed values. Time series fore-  
 53      casting plays a significant role in many application domains such as econometrics  
 54      (Lütkepohl et al., 2004), mathematical finance (Taylor, 2008), electroencephalography  
 55      (Blinowska & Malinowski, 1991), astronomy (Weigend, 2018) and communications en-  
 56      gineering (Brown, 2004). Due to the financial importance of large scale forecasting of  
 57      commodity values, time series forecasting has been led by disciplines associated with eco-  
 58      nomics. Economic time series forecasting has led to standard time series prediction tools  
 59      such as SARIMA (Box et al., 2015; Hamilton, 1994; Brockwell & Davis, 2016); a key fore-  
 60      casting tool evaluated within our work. More recently, Recurrent Neural Networks have  
 61      been effectively used for time series prediction using methods such as LSTMs (Hochreiter  
 62      & Schmidhuber, 1997; Greff et al., 2017) and sequence to sequence (Seq2Seq) methods (Sutskever  
 63      et al., 2014; Cho et al., 2014). LSTM and Seq2Seq methods are easily adapted to both

65 univariate or multivariate time series prediction (Rebane et al., 2018; Torres & Qiu, 2018).  
 66 It should be noted that some authors have identified weaknesses and limitations in us-  
 67 ing machine learning methods in general and LSTM methods in particular for time se-  
 68 ries forecasting e.g. (Chacón et al., 2020).

69 For many of the processes that contribute to InSAR measurements, we expect that  
 70 prior observations will not contain sufficient information to accurately predict future ob-  
 71 servations. This includes both signals of interest, such as sudden catastrophic failures,  
 72 and noise terms, such as turbulent atmospheric effects. However, some components of  
 73 the signal have repeating characteristics, such as multi-year trends and seasonal effects.  
 74 We begin by analysing the characteristics of the input dataset to select signals with re-  
 75 peating characteristics with a period of 1 year (section 3), and then focus on forecast-  
 76 ing over time periods of 1-9 months (section 4 and 5). Finally, we discuss the potential  
 77 applications and current limitations of time-series forecasting for Sentinel-1 InSAR data.

## 78 2 Case Study Dataset

### 79 2.1 InSAR Data

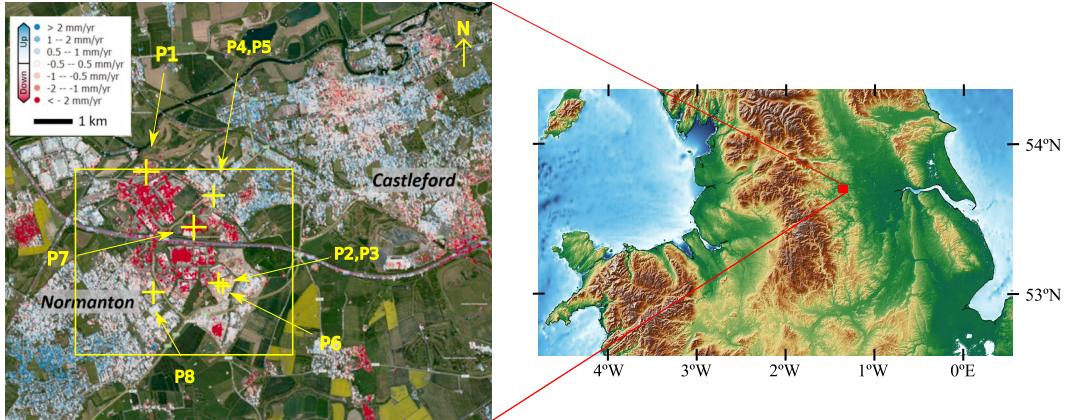
80 We test our algorithms on Sentinel-1 data processed by Satsense Ltd. using an algorithm  
 81 based on the RapidSAR approach (Spaans & Hooper, 2016). Atmospheric effects are the  
 82 dominant source of noise in most InSAR datasets and have been reduced within the Sat-  
 83 sense data through: (1) The removal of long wavelength signals from each InSAR im-  
 84 age using a Gaussian spatial filter. (2) The removal of short wavelength atmospheric sig-  
 85 nals using an APS (Atmospheric Phase Screen) filter. This isolates the random-in-time  
 86 effects using a highpass filter and then uses a low-pass spatial filter to estimate the spa-  
 87 tially correlated temporally random atmospheric effects. (3) Smoothing the displace-  
 88 ments in time using a per-time-series temporal filter to reduce the effects of overall temporal  
 89 noise which may include some residual atmospheric noise not removed by the APS fil-  
 90 ter.

91 Sentinel-1 acquires data every 6 days over Europe, but due to operational factors,  
 92 there are small infrequent gaps in the time-series, particularly in the first year when only  
 93 Sentinel-1A was operating. Since the algorithms proposed here require regularly sam-  
 94 pled data, we interpolate onto an even 6-day temporal grid as shown in Supplementary  
 95 Figure 1. Simple linear interpolation between neighbours is used to avoid unnecessary  
 96 assumptions.

### 97 2.2 Case Study Area

98 This project is part of the UK Digital Environment Programme and we use the subsi-  
 99 dence of the West Yorkshire coal mines as a case study (Burke et al., 2015; Lake et al.,  
 100 1992). Here we choose to work on the area around Normanton, which was mined until  
 101 the mid-1970s and where there is a high density of InSAR scatterers (Figure 1). The area  
 102 is currently subsiding at a rate of up to 15mm/yr and superimposed on this are seasonal  
 103 signals, particularly associated with some of the large warehouse buildings in the area.

104 A subset of the time series (points P1-P8) have been selected for further analysis  
 105 and forecasting experiments, and these are shown in Figure 1. P1, P3 and P8 illustrate  
 106 the combination of a (downward) trend and seasonality; P4-P6 have a strong seasonal  
 107 signal, but no long-term trend, and P7 and P8 show trends without seasonality. Points  
 108 P1-P6 were selected as being the top six seasonal signals according to the analysis in sec-  
 109 tion 3 and points P7 and P8 the lowest. P1-P3 and P6-P7 are car parks; P4 and P5 are  
 110 the roofs of a house and P8 is the roof of the XPO Logistics warehouse.



**Figure 1.** Large scale subsidence in West Yorkshire due to historical shallow coal mining. Central figure shows colour coded motion magnitudes. Points P1-P8 show the chosen points for analysis. P1-P3 illustrate the combination of a (downward) trend and seasonality; P4-P6 have a strong seasonal signal, but no long-term trend, and P7 shows a trend without seasonality. P1-P3 and P6-P7 are car parks; P4 and P5 are the roofs of a house and P8 is the roof of the XPO Logistics warehouse. The colour scale is saturated over 2mm/yr. Yellow inset square shows region analysed by Figure 2. The time series signals P1-P8 are shown in Figure 8.

### 3 Seasonal Signals in the InSAR Dataset

#### 3.1 Measures of Seasonality

Our hypothesis is that InSAR signals contain some periodic components, for which time series forecasting may be useful. For this application, we chose to focus on the most common natural periodic variations, those that occur annually. We start by testing the most commonly used method for estimating and removing seasonal components of geodetic time series, namely sinusoid fitting (Watson et al., 2002; Colesanti et al., 2003): see section 3.1.1. However, this measures the correlation with purely sinusoidal behaviour and could potentially exclude periodic signals with other non sinusoidal but repeating waveforms. A summary of an exhaustive variety of methods of detecting seasonality (Hartmann et al., 1992; Zubaidi et al., 2018; Hylleberg, 1995) is given in Supp. Table 1. and a detailed review of most suitable (to this application) methods is given in the remainder of this section. We focus on methods that are able to generate quantitative measures of annual seasonality rather than simple detection and can be used to analyse pre-defined periods (12 months) rather than estimate the period of seasonality. Based on these criteria, we select ‘Seasonal and Trend decomposition using Loess’ (STL: see section 3.1.2)(R. B. Cleveland et al., 1990) and the AutoCorrelation Function (ACF: see section 3.1.3)(Chen & Boccelli, 2018) for further study.

The choice of whether or not to normalise the seasonality measures is a key design decision. With normalisation, the amplitude of the seasonality will be disregarded, but if there is no normalisation, high amplitude stochastic signal components will often mask truly seasonal signals with small amplitude. For this reason, all three considered seasonality measures are normalised.

### 134 3.1.1 Sinusoid Fitting and Correlation (*Sin*) Method

135 We fit a sinusoid of fixed frequency (12 months) to the detrended time series using a least  
 136 squares method and extract the amplitude and phase parameters. An obvious measure  
 137 of seasonality is the magnitude of the fitted sinusoid, however, in this case, large mag-  
 138 nitude signals that are not particularly seasonal will produce a bigger seasonality index  
 139 than smaller magnitude signals that are truly seasonal. Instead, we define the seasonal  
 140 index for this method to be the normalised correlation between the training signal and  
 141 the fitted sinusoid,

$$142 \text{SIndex}_{\text{Sin}} = \rho(W_t, \hat{W}_{\text{sin}}), \quad (1)$$

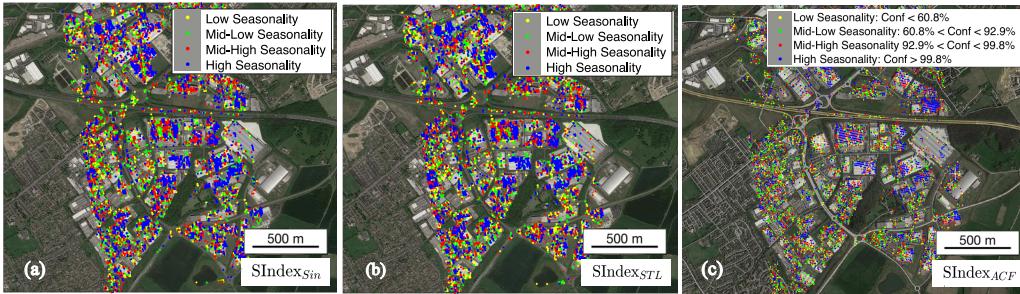
where  $\rho$  is normalised correlation and  $\hat{W}_{\text{sin}}$  is the fitted sinusoid.

### 143 3.1.2 STL decomposition

144 The concept of a “seasonal decomposition” of a time series signal means that the time  
 145 series can be decomposed into a sum (or a product) of three components: a trend (T),  
 146 a seasonal component (S), and a residual (R). We have used the common implementa-  
 147 tion of STL as initially described by Cleveland (R. B. Cleveland et al., 1990) assuming  
 148 an additive STL model. This implementation uses Loess smoothing, which uses itera-  
 149 tive sliding window regression to generate smooth functions (seasonal and trend)  
 150 (W. S. Cleveland, 1979). First, Loess smoothing is applied to remove the seasonal com-  
 151 ponent then a separate Loess smoothing is applied to remove the trend. The remaining  
 152 component is the residual.

153 A logical measure of the seasonality can then be defined using the ratio of the vari-  
 154 ance of the residual (R) to the variance of the signal without the trend (R+S). As this  
 155 ratio increases as seasonality decreases, we define seasonality as follows.  $\text{SIndex}_{\text{STL}}$  is  
 156 mathematically well behaved and mostly varies from 0 to 1.

$$157 \text{SIndex}_{\text{STL}} = 1.0 - \frac{\text{Var}[R]}{\text{Var}[R + S]} \quad (2)$$



**Figure 2.** Dataframe of InSAR datapoints in Normanton area grouped by levels of seasonality using: (a)  $\text{SIndex}_{\text{Sin}}$ , (b)  $\text{SIndex}_{\text{STL}}$ , and (c)  $\text{SIndex}_{\text{ACF}}$ . The  $\text{SIndex}_{\text{ACF}}$  sub figure is divided into four ranges of confidence bounds. Confidence is calculated as the rejection of the Null hypothesis that the ACF value is insignificant using the standard errors under the assumption of a Gaussian source (as used by the MATLAB `autocorr` function). Seasonality indices  $\text{SIndex}_{\text{STL}}$  and  $\text{SIndex}_{\text{Sin}}$  are divided into four equal and sorted ranges of seasonality indexed by colour.

### 157 3.1.3 Autocorrelation Function (ACF) Method

158 The autocorrelation function (ACF) measures how self-similar a signal is by measuring  
 159 the correlation of the signal with shifted versions of itself (Chen & Boccelli, 2018; Carla

160 et al., 2016). These shifts are known as lags and in this case, we are only interested in  
 161 the lag corresponding to 12 months. As the InSAR signal is sampled every 6 days (from  
 162 2015 to 2018) the lag is set to be 60. SIndex<sub>ACF</sub> is well behaved and varies from 1 (per-  
 163 fect correlation) to -1 (perfect anti-correlation). It is defined in (3) where  $\rho$  is the nor-  
 164 malised ACF function (with lag 60).

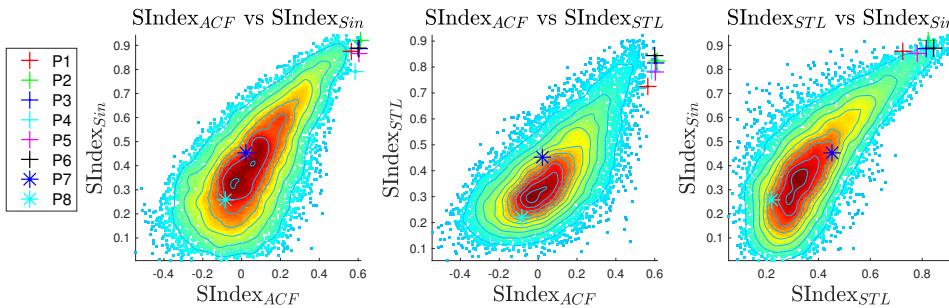
$$\text{SIndex}_{\text{ACF}} = \rho_{60}(W_t) \quad (3)$$

165 In order to properly estimate seasonality, isolated from the influence of trend, the trend  
 166 is removed by fitting a second degree polynomial to the InSAR time series and subtract-  
 167 ing it when using the ACF method. A second-degree polynomial was chosen to properly  
 168 model DC variations over the trained signal (this is not done for the STL method where  
 169 the trend is extracted independently). Confidence values can then be calculated as the  
 170 rejection of the null hypothesis that the ACF value is insignificant using standard errors  
 171 under the assumption of a Gaussian source (see Figure 2(a)).

### 172 3.2 Comparison of seasonality measures

173 For the ACF method (Figure 2(a)), seasonality correlates well with land use type, with  
 174 the highest values attributed to the roofs of particular buildings (for example the Wake-  
 175 field ASDA distribution centre). This suggests the dominant effect in such cases is ther-  
 176 mal expansion and contraction of man-made structures. Figures 2(b) and 2(c) show that  
 177 sinusoid fitting and STL methods are less spatially correlated (in terms of the different  
 178 seasonality magnitudes) when compared to the ACF based measure.

179 Figure 3 shows a comparison of the seasonality measures SIndex<sub>Sin</sub>, SIndex<sub>STL</sub> and  
 180 SIndex<sub>ACF</sub> for all the datapoints in Normanton region (with points P1-8 labelled). The  
 181 approximately linear relationship between the measures demonstrates that they are broadly  
 182 comparable, and the points P1-6 are classified as highly seasonal by all three indices, whereas  
 183 P7-8 lie with the majority of points which are not seasonal. However, there is consid-  
 184 erable scatter showing that the three indices identify different types of seasonality, with  
 185 especially large differences between the ACF and STL measures. We use the ACF mea-  
 186 sure for the subsequent experiments (due to the advantages given in table S1: supple-  
 187 mentary material and because it appears to be the most spatially coherent in Figure 2.).



**Figure 3.** Comparison of Seasonality Measures. SIndex<sub>Sin</sub> is the normalised correlation be-  
 between the signal and the best-fitting sinusoid. SIndex<sub>STL</sub> is based on a seasonal decomposi-  
 tion (STL) and defined as the ratio of the variance of the residual (R) to the variance of  
 the signal without the trend (R+S). SIndex<sub>ACF</sub> is the normalised autocorrelation function with a  
 period of 1 year (or 60 datapoints).

## 188 4 Ground Motion Forecasting

189 The task of forecasting InSAR time series can be approached in one of three ways: 1)  
 190 Future displacements forecast on each point individually, using only information from  
 191 that point (Mazzanti et al., 2011); 2) Future displacements can be forecast for each point  
 192 individually, using the time series itself and a selected group of related time series; 3)  
 193 Groups of time series can be forecast in a multidimensional sense: Seq2Seq models  
 194 (Rebane et al., 2018) and LSTM models (Torres & Qiu, 2018). For this proof-of-concept,  
 195 we have focused on the first two approaches for simplicity.

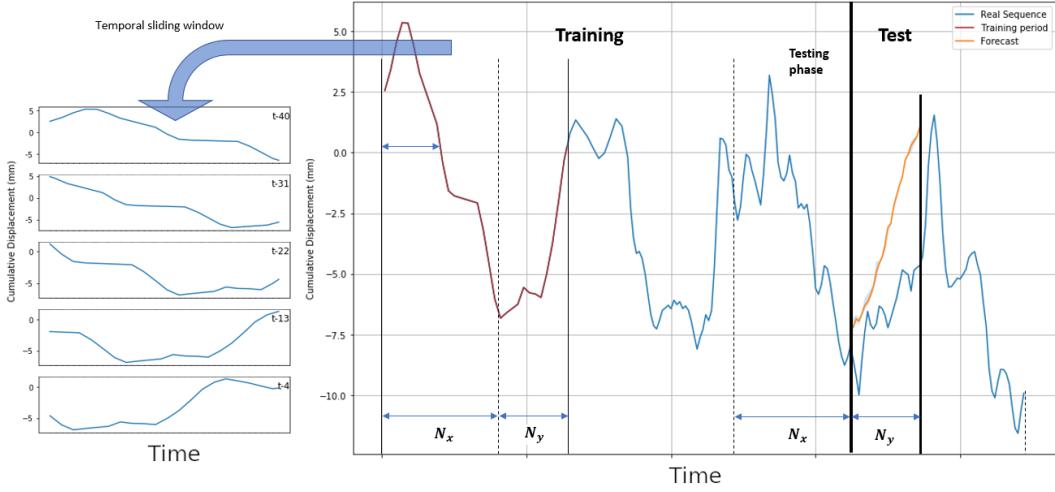
196 In this paper, we evaluate the forecasting methods by dividing each signal into a  
 197 “training” sub-range and a contiguous “testing” sub-range in order to be able to gen-  
 198 erate objective evaluations of each forecast. That is, given an entire temporal signal of  
 199  $T$  datapoints  $W = \{w_1, w_2, \dots, w_T\}$ , we define “today” as being timestep  $= t$  and our  
 200 goal is to predict a new signal  $\hat{y} = \{\hat{w}_{t+1}, \hat{w}_{t+2}, \dots, \hat{w}_{t+N_y}\}$  as similar as possible to the  
 201 original sub-signal  $y = \{w_{t+1}, w_{t+2}, \dots, w_{t+N_y}\} \in W$ . We also define the training set  
 202 as  $W_t = \{w_1, w_2, \dots, w_t\} \in W$ . The value  $N_y$  is a positive scalar integer which deter-  
 203 mines the time range to be forecast - i.e. the number of future observations. We set  $N_y$   
 204 to approximately 9 months (264 days) for all experiments, but evaluate the prediction  
 205 over time periods of 1-9 months. The value  $N_x$  is a positive scalar integer which deter-  
 206 mines the period of time for training i.e. the number of past observations (in samples  
 207 where samples are every 6 days). An illustrative example of the predicted and test sig-  
 208 nals is shown on the right of Figure 4.

209 A summary of all the forecast methods compared is given in Table 1. In order to  
 210 compare the SARIMA and LSTM approaches with previously used methods, we include  
 211 a standard sinusoid fitting algorithm (Watson et al., 2002), and project the fit forward  
 212 in time. A sinusoid and trend are fitted to the same part of the each of the time series  
 213 as used for training the other methods (i.e.  $W_t$ ) and future values extrapolated using the  
 214 resulting parameterisation.

### 215 4.1 Long Short-Term Memory (LSTM) Networks

216 A Long Short-Term Memory (LSTM) (Hochreiter & Schmidhuber, 1997; Greff et al., 2017)  
 217 network is a Recurrent Neural Network (RNN) (Rumelhart et al., 1986) architecture used  
 218 in the field of deep learning for time series data. LSTMs keep track of arbitrary long-  
 219 term dependencies in the input sequences, and they can scale to much longer sequences  
 220 than conventional RNN networks. They are designed to process sequences of variable  
 221 lengths, where parameters are shared with all previous output members. LSTMs have  
 222 the ability to add or remove information to a temporal learning “state”. This is care-  
 223 fully regulated by structures called gates. The learning selectively *keeps* some part of the  
 224 past (using the temporal states) and *forgets* others (using “forget” gates). LSTMs are  
 225 commonly used for classification applications.

226 The use of LSTMs for time series forecasting of InSAR derived ground motion data  
 227 has both limitations and advantages. Due to the capabilities of LSTMs to model long  
 228 and short term patterns, LSTMs should be able to flexibly model and predict time se-  
 229 ries data (including annual, seasonal and trend variations) and are easily extendible to  
 230 larger and multivariate datasets. Furthermore, the use of a supervised learning system  
 231 such as an LSTM based system also has the advantage of modelling any predictable vari-  
 232 ations found in the data not just those defined in ad-hoc methods such as SARIMA or  
 233 Sinusoid prediction. However, the sliding window approach required by any supervised  
 234 machine learning method is a significant limitation as it reduces the amount of training  
 235 data but also limits the internal LSTM state explicitly modelling very long periodic sig-  
 236 nals components. The fact that LSTMs are not able specifically to analyse a single sea-  
 237 sonality period of interest (in this case annual seasonality) is both an advantage and a  
 238 disadvantage. In the case of our data, figure S7 shows that there is no distinct frequency



**Figure 4.** Example of a multi-step approach for time-series forecasting. The considered signal is split into training and test sets. The training set contains the observable data and the test signal represents the future observations in the test set. This test signal is to be compared with the predicted signal obtained by our method as well as its associated prediction error. Our multi-step approach for LSTMs reframes the whole training data into temporal sliding windows of sizes  $N_x$  and  $N_y$  for past and future observations, respectively.

peak. However, it may be useful to focus on annual seasonality due to physical process causing cyclic variations over annual periods. Finally, supervised learning approaches such as LSTMs are very dependent on the amount of data. There are only 91 training data points for a sliding window in a single signal we have used (LSTM1).

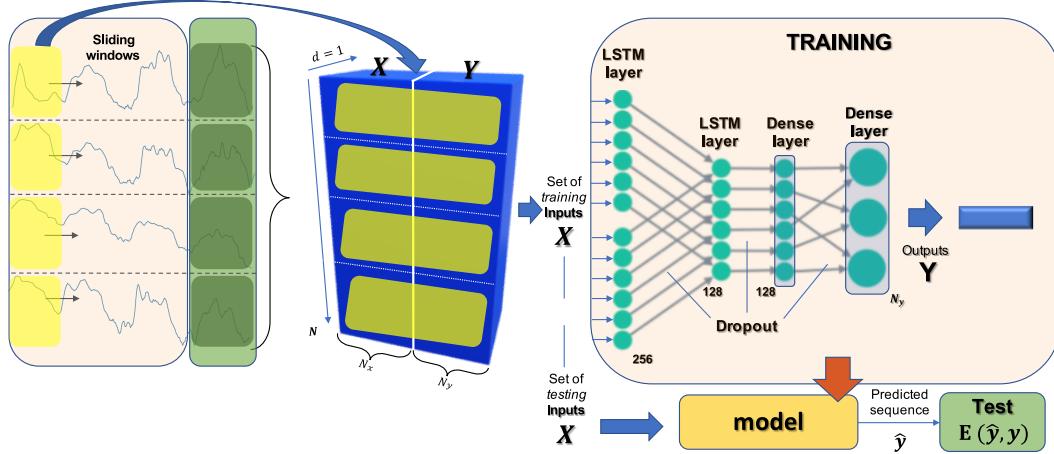
Within this application we are using LSTM within the supervised regression framework illustrated in Figure 4 where the output of the network  $\hat{y}$  is an array of length  $N_y$ . The multi-variate and multi-signal methods described below are significantly different. The multi-variate (LSTM4) method has a multi-dimensional feature input into the network whereas the multi-signal methods are still one dimensional but have more than one input signal (LSTM2-3).

#### 4.1.1 Univariate LSTM: LSTM1

Figure 4 shows the univariate case of forecasting ground motion using a supervised LSTM network. A sliding window forms a training data frame (for a single signal) of inputs ( $\mathbf{X}$ ) and outputs ( $\mathbf{Y}$ ) to train the network. Once trained, the testing input ( $\mathbf{X}_{\text{test}}$ ) is ingested into the network to generate the forecast  $\hat{y}$  approximating the true sequence  $y$ . This requires a final layer in the network to generate a vector the same length as  $y$ . This is done using a fully connected dense layer without any subsequent pooling as illustrated in Figure 5. For the LSTM experiments (resulting in figures 6-9)  $N_x$  and  $N_y$  are set to 9 months (44 samples equating to 264 days). This is considered to be long enough to characterise the seasonal nature of the signals, be able quickly to adapt to changes and also have the maximum amount of training data from the sliding window. This design choice is evaluated in section 5.1.1.. We use a network of two LSTM layers fully connected to a dense layer outputting the  $N_y$  regression outputs. The first layer has a large number of nodes in order to effectively characterise the data across all models (for simplicity and communicability). Large numbers of nodes and complex network architectures tend to overfit. However, we have chosen to deal with overfitting using dropout layers. Each layer

265 has an integrated dropout function (set to a dropout factor of 0.5). The optimisation  
 266 was based on the ADAM method (Kingma & Ba, 2015) and Mean Square Error (MSE)  
 267 as the loss function. We train our networks using 2000 iterations (epochs) to achieve con-  
 268 vergence. The number of samples used in each method LSTM1-4 is given in Table-1.

#### 269 4.1.2 Multi-Signal LSTM: LSTM2-4



270 **Figure 5. Multi-Signal LSTM.** Every signal in the training set is split into training and  
 271 test sets in the multi-signal approach. First, we use the training sets to frame every signal as a  
 272 supervised machine learning problem, constructed by a set of inputs  $\mathbf{X}$  and outputs  $\mathbf{Y}$ . Each  
 273 considered time-step for the sliding window for each signal becomes a sample in the feature space  
 274 of input  $\mathbf{X}$  and output  $\mathbf{Y}$  of the network.

275 We adapt the univariate approach shown in Figure 4 to include data from a set of train-  
 276 ing signals (described below). This multi-signal LSTM is illustrated in Figure 5. This  
 277 system uses the same network structure as above but vertically concatenates all of the  
 278 sliding window data from a set of training signals (described below). The testing data  
 279 remains the same (i.e. a single spatial sample in each case). The LSTM2 system uses the  
 280 top six seasonal signals for training. The LSTM3 system uses the top 1% of the seasonal  
 281 signals (using the SIndex<sub>ACF</sub> method) for training. Conversely, LSTM4 uses the eight  
 282 spatially closest time series signals as features in an eight-dimensional multivariate LSTM  
 283 input. A multivariate LSTM architecture is then used to generate a univariate forecast  
 284 from the multivariate InSAR derived ground motion time series data. The size of the  
 285 associated input and output network tensors are indicated in Table 1.

#### 286 4.1.3 Seq2Seq LSTMs: Seq2Seq1-4

287 Sequence to sequence (Seq2Seq) is an encoder-decoder deep learning architecture for mak-  
 288 ing multi-step predictions (Sutskever et al., 2014; Cho et al., 2014). The previous meth-  
 289 ods (LSTM1-4) generated the prediction vector using the single output of an LSTM layer  
 290 together with dense and fully connected layers (with a final vector regression output).  
 291 Seq2Seq methods have an independent encoder that analyses the input time sequence  
 292 and generates a characterising set of states that are subsequently input into the decoder.  
 293 We have used a single LSTM layer as the encoder that outputs the LSTM states of the  
 294 input time series data as an initial stage. These output states are then copied multiple

times (with the number of copies being the required length of the prediction vector output). These copies then form a multidimensional time series input to a decoder (another single LSTM layer). The time distributed outputs are then input into time distributed dense layers outputting a vector forecast result  $\hat{y}$ . Each method LSTM1-4 has been modified to include a Seq2Seq architecture to form methods Seq2Seq1-4 respectively i.e. the other architectural forms and input/output data structures are equivalent for these two sets of methods.

## 4.2 SARIMA

SARIMA is an analysis and prediction method for time series data (Box et al., 2015; Hamilton, 1994; Brockwell & Davis, 2016). It is used to model non-stationary data series, where the data are not statistically consistent across time e.g. mean and variance varies with time. It is often combined with Kalman filters that use the model to predict future values of a time series signal. SARIMA is an analysis tool primarily used to model economic data and is able to identify, model and predict both trend and seasonality (and their variations) over time. SARIMA consists of two sets of forecasting models: trend and seasonality. Each of these two models are divided into three submodels: an autoregressive model (AR) and a Moving Average (MA) model in order to model time variations (“tendencies”). The MA model is the equivalent of an estimated Finite Impulse Response (FIR) filter that just weights recent inputs to combine into an estimated output. Conversely, the AR model is an estimated all-pole or Infinite Impulse Response Filter (IIR) that uses a feedback loop to estimate output given a weighted sum of previous outputs. The input is often further locally differenced (the I stage) to model changes in offset (the third submodel).

The model is comprised of these three sub-models (AR, MA and I) estimated directly on the data to model trend but also over a set lag directly related to the seasonality of the signal. A SARIMA model is then defined as the order of these six models (plus the analysis seasonality lag  $m$ ):

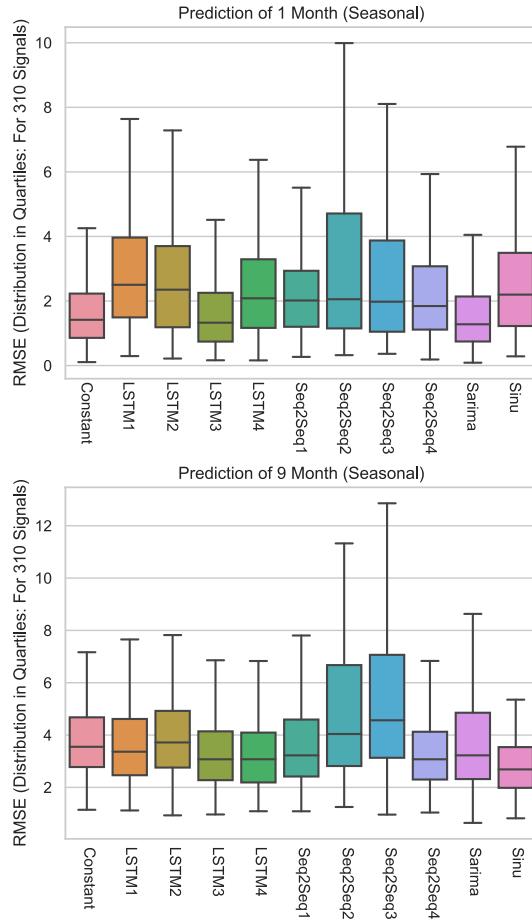
$$\text{SARIMA} \quad \underbrace{(p, d, q)}_{\substack{\uparrow \\ \text{Non-Seasonal Part} \\ \text{of the Model}}} \quad \underbrace{(P, D, Q)_m}_{\substack{\uparrow \\ \text{Seasonal Part} \\ \text{of the Model}}}$$

where  $p$  is the order of the AR term,  $q$  is the order of the MA term,  $d$  is the number of differencing operations required to make the time series stationary,  $P$  is the order of the AR seasonality term,  $Q$  is the order of the MA seasonality term,  $D$  is the number of differencing operations required to make the seasonal time series stationary and  $m$  is the seasonality lag.

The parameters of the SARIMA model are commonly not estimated automatically i.e. the statistics and correlation of the time series signal are analysed by hand and the parameters are tuned until the signal (when compensated by the found parameters) is considered to be stationary. However, recent automatic parameter estimation methods do a minimisation search on some training data to determine the best combination of SARIMA parameters (Hyndman et al., 2007). This method estimates the stationarity of the signal under the parameters and specifically uses the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) estimators to compare models. The lower these values, the better the model fits the data (Hyndman et al., 2007).

Here, SARIMA parameters are fitted to the training data using Hyndman’s method (Hyndman et al., 2007). A typical model for the analysed InSAR time series below was SARIMA(3, 0, 2)(1, 1, 0)<sub>60</sub>. The parameters are estimated using the same part of each

335 of the time series as used for training with LSTMs (i.e.  $W_t$ ) and then SARIMA is used  
 336 to predict the same part of each time series as with LSTMs (i.e.  $W_y$ ). Kalman filters are  
 337 used with the SARIMA model in standard SARIMA implementation packages to pre-  
 338 dict future time series values. We have used such an implementation for time series fore-  
 339 casting.

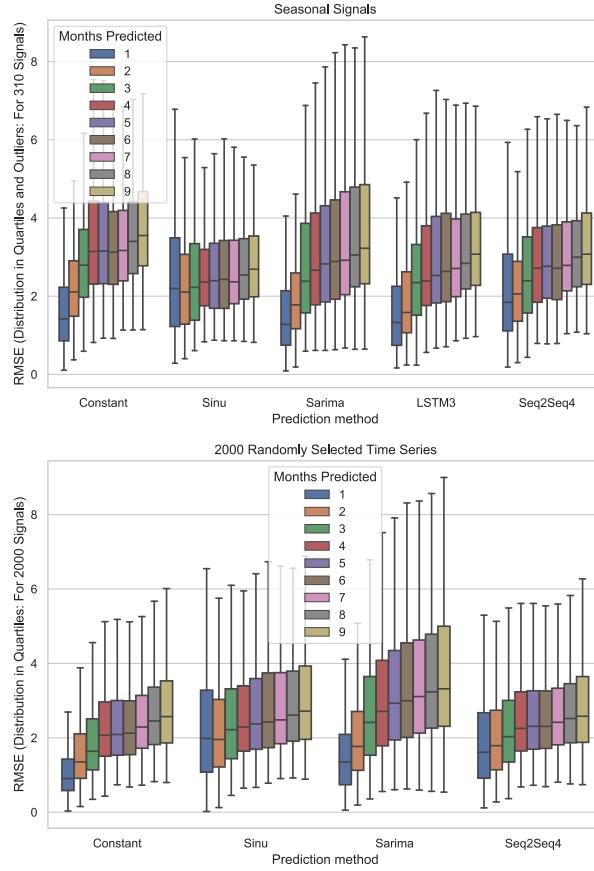


340 **Figure 6.** Forecast performance on seasonal signals. The boxplot lines represent the quartiles  
 341 of the RMSE distribution for all 310 signals with coloured area being the central two quartiles  
 342 and the central line being the median. a) performance over 1 month; b) performance over 9  
 343 months.  
 344

## 345 5 Forecast Performance

### 346 5.1 Seasonal Signals

347 We test the forecasting performance of LSTMs and SARIMA on a set of 310 highly sea-  
 348 sonal signals selected using the SIndex<sub>ACF</sub> metric. We benchmark the results against si-  
 349 nusoid extrapolation and a constant value prediction. To assess the performance of each  
 350 model, we use the Root Mean Square Error,  $RMSE(\hat{y}) = \sqrt{E((\hat{y} - y)^2)}$  (Figures 6, 7).  
 351 We also consider normalised RMSE and define n1RMSE and n2RMSE as the RMSE of  
 352 the prediction normalised against the variance and constant value prediction respectively  
 353 (Supp. Figures 2-5). The RMSE distributions are displayed in the form of a boxplot that  
 354

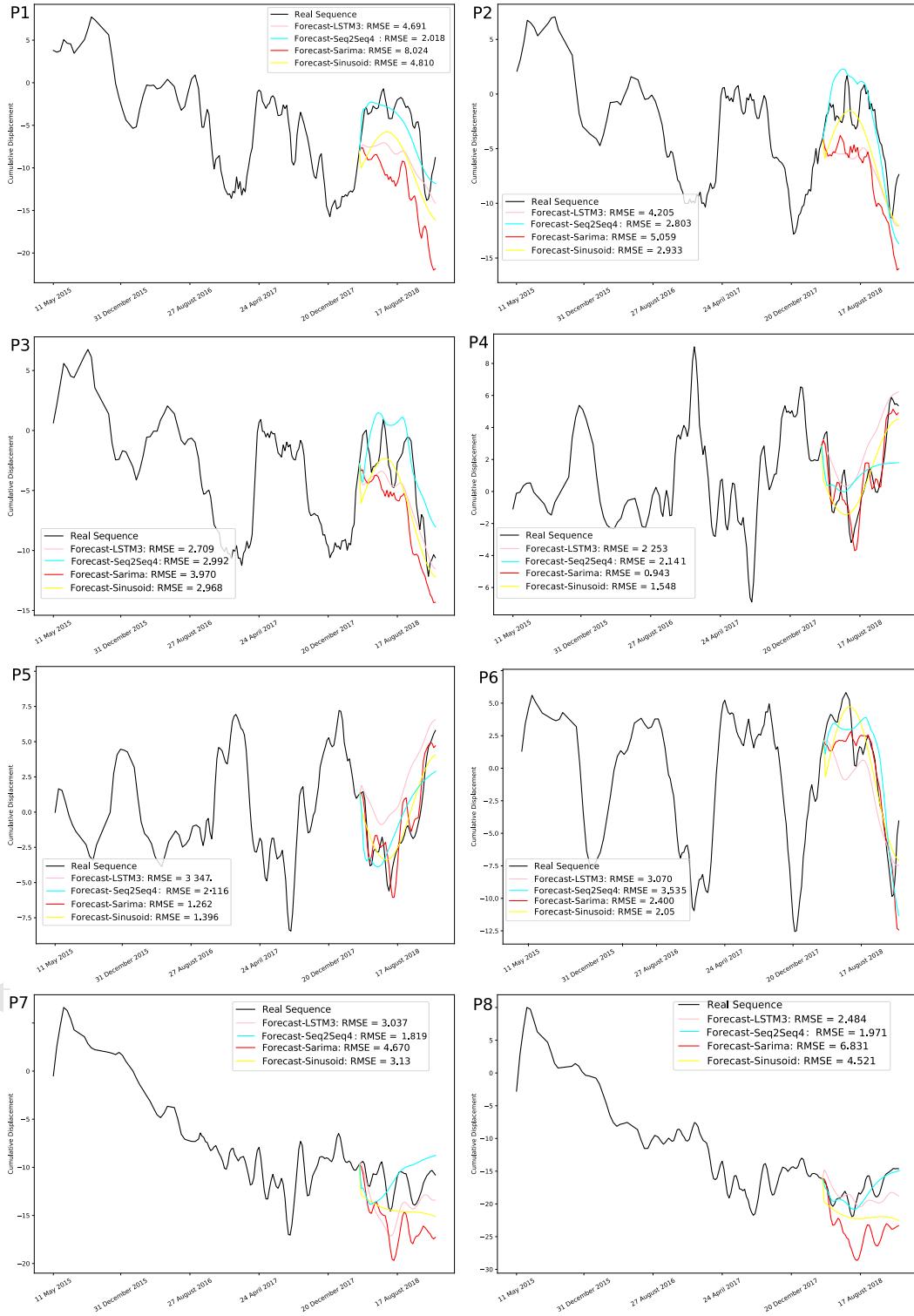


**Figure 7.** Forecast performance according to prediction window for a) 310 seasonal signals and b) 2000 randomly-selected signals. The boxplot lines represent the quartiles of the RMSE distribution with coloured area being the central two quartiles and the central line being the median.

349 includes the quartiles of the distribution (the middle line in each box is the distribution  
 350 median).

351 For a one month prediction (Figure 6a), the best performing methods were SARIMA  
 352 and LSTM3, which had a slightly lower median RMSE than the constant value predic-  
 353 tion. Of the Seq2Seq methods, the lowest median RMSE was produced by the univari-  
 354 ate version Seq2Seq1 and Seq2Seq4, which was trained with using the 8 geographically  
 355 closest points. For these short time periods, the sinusoidal extrapolation method (Sinu)  
 356 had a median RMSE value considerably higher than that of the constant value pre-  
 357 diction. Conversely, for longer time periods (Figure 6b), the lowest median RMSE was si-  
 358 nusoid extrapolation, with a median RMSE value about 75% of the constant value pre-  
 359 diction. Of the LSTMs, the lowest median RSME values were associated with LSTM3  
 360 and LSTM4, while Seq2Seq1 and Seq2Seq4 had lower median RMSE values than the multi-  
 361 signal Seq2Seq methods (Seq2Seq2-3). Over these time periods, most of the methods had  
 362 lower median RMSE values than the constant value prediction, with only LSTM2, Seq2Seq2  
 363 and Seq2Seq3 performing worse (when considering median value of n2RMSE: Supp. Fig-  
 364 ure 3).

Accepted Article



**Figure 8.** InSAR example forecasts of eight sample signals: LSTMs, SARIMA and Sinusoid Fitting.  $N_y = 9$  months ( $N_x = 9$  months for LSTMs). The top six signals are highly seasonal signals whereas the bottom two are highly un-seasonal

365 For all the time periods considered, the multi-signal Seq2Seq models (Seq2Seq2-  
 366 3) trained using a set of seasonal signals had a lower median RMSE than the univari-  
 367 ate case (Seq2Seq1). We conclude that any improvements gained by having a larger train-  
 368 ing dataset are offset by the potentially spatially unrelated data statistics and charac-  
 369 teristics. However, Seq2Seq4, which was trained using geographically close signals, had  
 370 a median RMSE value a little better than the univariate case (Seq2Seq1) suggesting that  
 371 geographically close points have more similar signals, as for example, they may be lo-  
 372 cated on the same structure.

373 Based on this assessment, we select LSTM3, Seq2Seq4 and SARIMA for further  
 374 analysis and some examples of the predicted and real time series are shown in Figure 8.  
 375 Points P1-P6 were selected as they have the most seasonal characteristics signals as de-  
 376 fined by  $SIndex_{ACF}$ . All methods capture some aspects of the signal, and the time se-  
 377 ries plots are helpful in identifying sources of misfit. For example, sinusoid extrapola-  
 378 tion is a global fitting method, so there is often a discontinuity between the training and  
 379 prediction data (e.g. P3, P6; Figure 8), which explains why the RMSE is high when short  
 380 prediction periods are considered (Figure 6a). Similarly, the SARIMA results can be seen  
 381 to characterise the sub-seasonal variations of many of the example 6 signals, but for P1  
 382 and P2, the trend has not been accurately estimated and the prediction, although plau-  
 383 sible in shape, has an inaccurate offset.

384 The results in Figure 6 suggests that performance varies according to prediction  
 385 window, so we test the selected methods over periods of 1-9 months and compare the  
 386 distribution of RSME values (Figure 7). The lowest RMSE values are obtained for SARIMA  
 387 when considering short term predictions of < 3 months, whereas sinusoid extrapolation  
 388 performs best for predictions of > 6 months. As expected RMSE increases with increas-  
 389 ing prediction window: the constant value prediction has a median RMSE value of 1.4  
 390 cm for a 1 month window, increasing to 3.6 cm for a 9 month window. Normalising the  
 391 RMSE to the RMSE value of the constant value prediction (n2RMSE, Supp. Figure 5)  
 392 removes this effect, and shows that SARIMA and Seq2Seq4 outperform the constant value  
 393 prediction for all windows, whereas Sinu and Seq2Seq4 only perform better when fore-  
 394 casting 3 or more months into the future.

395 The multi-signal LSTM (LSTM3) gave the best results for short term prediction  
 396 (< 3 months). SARIMA also gave good results for short term prediction but gave sig-  
 397 nificantly worse results (compared to LSTM3) for predicting many months into the fu-  
 398 ture (see table 1 for a detailed description of the different training methods associated  
 399 with each method). The performance of the Seq2Seq4 method was virtually identical to  
 400 the LSTM3 method for a period of 9 months but had a slightly larger median error (by  
 401 0.3cm) for 1 month.

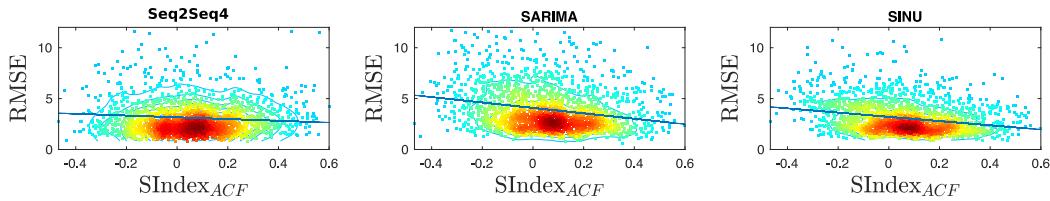
#### 402 ***5.1.1 Analysis of Retrained Models Varying $N_x$ and $N_y$***

403 An independent set of experiments were conducted to evaluate the effect of changing the  
 404 lengths of  $N_x$  and  $N_y$  on the RMSE performance for an individually retrained machine  
 405 learning system (specifically Seq2Seq1). These experiments (see Figure S7) show that  
 406 varying varying  $N_x$  either makes no difference (for  $N_x=18$  months compared to  $N_x=9$   
 407 months) or actually decreases performance ( $N_x = 27$  months compared to  $N_x=18$  or  $N_x=9$   
 408 months). We attribute this to be because of the sliding window approach where a wide  
 409 range of phases of the input data are modelled across the whole temporal range of data.  
 410 When large values of  $N_x$  are chosen, the amount of training data decreases resulting in  
 411 the decrease in performance for  $N_x=27$ . Results from these experiments are shown in  
 412 Figure S7. Although the results in figure 6 and 7 are generated from fixing  $N_y=9$  and  
 413 evaluating the performance over different. The results shown in Figure S7 show that there  
 414 is very little difference in retraining the models for varying values of  $N_y$ .

## 415 5.2 Randomly Selected Signals

416 Finally, we select 2000 points at random from the Normanton dataset with no regard  
 417 to seasonality and test the methods that performed best on the seasonal signals: SARIMA  
 418 and Seq2Seq4. Descriptions of training methods for SARIMA and Seq2Seq4 are given  
 419 in sections 4.2 and 4.1.3 respectively. We no longer consider LSTM3 since it was trained  
 420 specifically on highly seasonal signals. Points P7-P8 shown in Figure 8 illustrate the chal-  
 421 lenges of time-series prediction for non-seasonal signals. Figure 7 shows that the rela-  
 422 tive variation in RMSE with prediction window is similar to that for seasonal signals.  
 423 However, this figure shows that none of the methods perform better than the constant  
 424 value prediction when signals are randomly selected (see also Supp. Figure 5).

425 Figure 9 shows the relationship between forecast performance (RMSE) and season-  
 426 ality ( $SIndex_{ACF}$ ) for a prediction window of 9 months for the Seq2Seq4, SARIMA and  
 427 Sinu Methods. For Seq2Seq4, the forecast performance appears independent of season-  
 428 ality, whereas the SARIMA and Sinu methods perform better (decreased RMSE) with  
 429 increased seasonality. To test the statistical significance we calculate Pearson's corre-  
 430 lation coefficient (and associated p-value) for each relationship illustrated in figure 9. The  
 431 Pearson correlation coefficients were: Seq2Seq4(-0.072), SARIMA(-0.223) and Sinu (-0.255).  
 432 The corresponding p-values for these correlation values were: Seq2Seq4(0.0014), SARIMA( $7.6 \times$   
 433  $10^{-24}$ ) and Sinu ( $3.9 \times 10^{-31}$ ). These values indicate that all three results are negatively  
 434 correlated. However, the Seq2Seq method is less negatively correlated than the other two  
 435 methods, but there is an extremely small confidence in the negative correlation of SARIMA  
 436 and Sinu (given by very small p-values). A similar pattern is seen for all prediction win-  
 437 dows (Supp. Figure 6).



438 **Figure 9.** RMSE vs Seasonality Seq2Seq4, SARIMA and Sinu Methods. Plots show are for  
 439 prediction windows of 9 months, with full results for predictions windows of 1-9 months shown in  
 440 Supp. Figure 6.

## 438 6 Discussion

439 Previous studies have reported annual variations in InSAR data associated with processes  
 440 such as tropospheric water vapour (Heleno et al., 2010), thermal contraction and expan-  
 441 sion (Lazecky et al., 2016), ground water (Bell et al., 2008) and freeze-thaw cycles (“Large-  
 442 scale InSAR monitoring of permafrost freeze-thaw cycles on the Tibetan Plateau”, n.d.).  
 443 We find that our dataset from the Normanton area of the United Kingdom also contains  
 444 signals with periodic variations, the strongest of which are clustered on large warehouses  
 445 suggesting the dominant effect here is thermal expansion and contraction of man-made  
 446 structures.

447 We test the ability of a range of established time series prediction methods to fore-  
 448 cast InSAR time series and find that several methods perform better than a constant  
 449 value prediction when signals dominated by periodic variations are considered. The low-  
 450 est median RMSE values are obtained for SARIMA when considering short term pre-  
 451 dictions (<3 months), whereas sinusoid extrapolation performs best for longer predic-  
 452 tions (>6 months). However, for non-seasonal signals, the simple extrapolation of a con-

**Table 1.** Summary of Forecasting Methods. For all methods  $N_x = 9$  months,  $N_y = 9$  months (i.e. 44 samples or 264 days)

Method	Definition
<b>Constant</b>	A constant value prediction, taking the last value of the training time series and extrapolating it for the whole of the test series.
<b>Sinu</b>	Sinusoid fitting method: A simplex gradient descent method was used to fit the amplitude and phase of a sinusoid to the data (together with the slope of a linear trend term).
<b>SARIMA</b>	SARIMA and Kalman Filter based prediction with parameters obtained using auto sarima (Hyndman et al., 2007).
<b>LSTM1</b>	Single signal used for prediction (based on the univariate method illustrated in Figure 4). Architecture included: two LSTM layers (first with 256 nodes and second with 128 nodes). The final state output of the second LSTM layer is connected to a dense layer of 128 nodes and then subsequently connected to an output layer with $N_y$ nodes. Dropout of level 0.5 is included between each layer, the activation function was ReLU, the loss was MSE, the optimiser was ADAM. The number of samples in the training sliding window is 91. 91 refers to the number of input datapoints in the sliding window. The training part of the signal is 179 samples, the sliding window is 88 samples ( $44:N_x$ , $44:N_y$ ) therefore the total number of sliding positions is 91. Size( $X$ ): (None $\times 44 \times 1$ ), Size( $Y$ ): (None $\times 44 \times 1$ ).
<b>LSTM2</b>	The six most seasonal signals (seasonality measured using SIndex <sub>ACF</sub> ) concatenated and used for training (Figure 5). The remaining architectural features for this as per LSTM1. The number of samples in the six sliding windows is $6 \times 91 = 546$ . Size( $X$ ): (None $\times 44 \times 1$ ), Size( $Y$ ): (None $\times 44 \times 1$ ).
<b>LSTM3</b>	The top 1% of the seasonal signals (seasonality measured using SIndex <sub>ACF</sub> ) concatenated and used for training as per LSTM2. The number of samples in the training sliding windows of the top 1% seasonal signal is 10738. Size( $X$ ): (None $\times 44 \times 1$ ), Size( $Y$ ): (None $\times 44 \times 1$ ).
<b>LSTM4</b>	The eight spatially closest time series signals (see Supp. Figure 1b) are formed into different features in the multivariate learning process (with a single dimensional feature predicted for the considered time series). The number of multivariate samples in the training sliding window is 91. Size( $X$ ): (None $\times 44 \times 8$ ), Size( $Y$ ): (None $\times 44 \times 1$ ).
<b>Seq2Seq1</b>	Seq2Seq architecture. The encoder was a single encoder LSTM layer (with 200 nodes) whose output was copied $N_y$ times. This time distributed output was then input into the decoder; a single LSTM layer (with 200 nodes). This was then input to a fully connected dense layer with a final single (but time distributed output layer) node. No dropout was used. The remaining aspects of this architecture were as per LSTM1. Variation of network parameters (within reasonable limits) did not change performance. A similar set of network parameters was chosen for simplicity of definition.
<b>Seq2Seq2</b>	Same as LSTM2 but with Seq2Seq architecture as described for Seq2Seq1.
<b>Seq2Seq3</b>	Same as LSTM3 but with Seq2Seq architecture as described for Seq2Seq1..
<b>Seq2Seq4</b>	Same as LSTM4 but with Seq2Seq architecture as described for Seq2Seq1.

stant function perform better overall than any of the more sophisticated time series prediction methods. Comparisons between seasonality and RMSE show a reduction in RMSE with increasing seasonality. However, this is shown to be statistically insignificant (given the very low confidence values in the Pearson correlation coefficients between RMSE and ACF based seasonality: SIndex<sub>ACF</sub>).

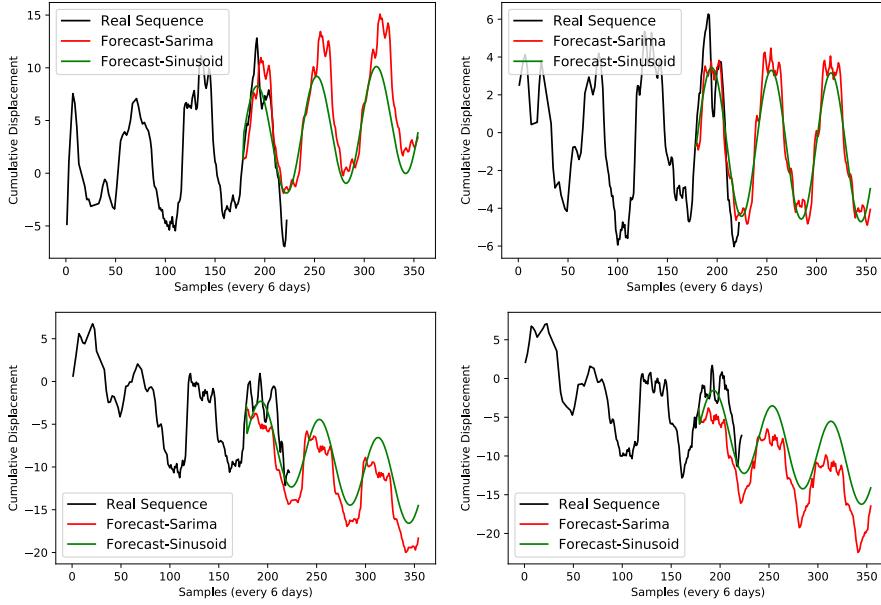
Much of the comparisons above only compare the median RMSE values, but the breadth of the distribution and scatter of outliers shows that the misfit is highly variable between observation points, even when seasonality is taken into account. Thus, if a prediction for a single observation point is required, there may be a large misfit in the prediction even for the best performing approaches, but a poorly performing method might produce a very accurate prediction.

The machine learning methods (LSTM and Seq2Seq) tested performed well in some cases, with the use of multivariate or concatenated signals improving the performance. However, it is interesting that they performed less well overall than simple extrapolation of a constant value (for non-seasonal signals) or a sinusoid (for seasonal signals). Interestingly, the performance of the machine learning methods only improved slightly with increasing seasonality, suggesting that they are failing to capture the periodic component of the signal, perhaps because they are only trained over 9 months. Poor performance in predicting financial time series using LSTMs has also been reported (Sirignano & Cont, 2019). This is assumed to be related to the non-stationary nature of the data and the inability of LSTMs to model feedback effectively. Improvements in prediction using LSTMs should follow through both large increases in training data (number of data sequences and length of sequences) together with the integration of SARIMA type feedback modelling.

In this study, we have focused on predictions for windows of less than the period of the signal (1 year), but both SARIMA and sinusoid extrapolation are able to predict for an arbitrary amount of time into the future. Figure 10 demonstrates that predictions for several years into the future show plausible time series, but unfortunately, no quantitative evaluation is possible until a longer dataset of measurements is acquired. Similarly, LSTM methods require a training window that is at least as long as the prediction window, and will require longer time-series before long-term predictions can be tested.

Real-time monitoring and ground motion forecasting of periodic signals from InSAR data could be used in one of two ways. The first of these is to predict seasonally varying ground motion signals that could otherwise obscure subtle deformation changes that could be precursors to rapid and critical collapses (Selvakumaran et al., 2018). In this case, the reduction in background noise could enable the detection of anomalous or unexpected behaviour. Alternatively, the periodic motion itself is of interest to insurance companies looking to forecast claims due to ground cracking and subsidence (Crilly, 2001), or bridge motion (Lazecky et al., 2016). The broad distribution of misfit values suggests these approaches will only be useful when considering the distribution of a large number of datapoints, and the probability of a good prediction for any single observation point is quite small.

This is a proof-of-concept study and the methods described here can be further refined. Possible future directions include testing different neural network architectures including convolutional LSTMs and attention based systems; the combination of SARIMA and LSTMs; the integration of spatial analysis using CNNs and multivariate prediction using Vector Autoregression. Future developments in machine learning and artificial intelligence may improve performance, but the lack of periodic or repeating signals within the dataset may always be a barrier to time series prediction.



**Figure 10.** InSAR example forecasts of seasonal signals (far into the future: 3 years)

### 502      6.1 LSTMs and Supervised Learning Methods for InSAR Ground Mo- 503      tion Forecasting

504      Given the stringent analysis of LSTM and supervised learning methods described in sec-  
505      tion 4.1, we believe the use and analysis of such methods justified. Despite the fact they  
506      did not give significantly better results than other methods in our results they should  
507      be considered as an effective model in situations where: 1) there is a larger amount of  
508      training data, 2) the signals are easier to predict, 3) there is a complex mixture of sea-  
509      sonality within the data and 4) the data is comprised of very high dimensional multi-  
510      variate signals (which would be difficult to analyse using the other methods discussed  
511      within this work). In our opinion, supervised machine learning methods such as LSTM  
512      methods should not be discounted by researchers but included in a toolbox of forecast-  
513      ing methods mapped out within this paper.

## 514      7 Conclusion

515      In this proof-of-concept study, we have tested a range of time series prediction tools on  
516      ground motion data collected using InSAR. For randomly-selected data, a simple con-  
517      stant value prediction outperforms both conventional time series analysis and forecast-  
518      ing methods such as SARIMA and supervised machine learning approaches such as LSTMs.  
519      This reflects the stochastic nature of the signals and the difficulties in using any trained  
520      system to predict far into the future. The time series prediction methods performed bet-  
521      ter on signals containing strong annual variations, and both LSTM based architectures  
522      and SARIMA performed better over short periods of time (less than three months) than  
523      the extrapolation of a sinusoidal function. Figure S7 (in the supplementary material) shows  
524      that there is no distinct frequency peak within the data and therefore. This suggests that  
525      there is not a strong temporal signal within the data but may go some way to explain  
526      the low performance of the temporal models.. This suggests that a pre-processing step  
527      could be used to select signals that are suitable for forecasting. However, further devel-  
528      opments in machine learning and artificial intelligence will be needed before time series  
529      predictions of InSAR data are sufficiently reliable to be used in practice.

## 530 8 Acknowledgements

531 We thank SatSense Ltd for access to their dataset over the Normanton/Castleford area.  
 532 This work was funded by the Digital Environment Programme under NE/S016104/1.  
 533 Data supporting this research are available in <https://www.satsense.com/data-portal/>.  
 534 The data is available to any user that signs an NDA and license agreement with Satsense.  
 535 The authors are not aware of any reason why the data would not be made available. Contact  
 536 Satsense at <https://www.satsense.com/data-portal/> to gain access to this data.

## 537 References

- 538 Aldiss, D., Burke, H., Chacksfield, B., Bingley, R., Teferle, N., Williams, S., ...  
 539 Press, N. (2014). Geological interpretation of current subsidence and uplift  
 540 in the london area, uk, as shown by high precision satellite-based surveying.  
 541 *Proceedings of the Geologists' Association*, 125(1), 1–13.
- 542 Anantrasirichai, N., Biggs, J., Albino, F., & Bull, D. (2019a). The application of  
 543 convolutional neural networks to detect slow, sustained deformation in insar  
 544 time series. *Geophysical Research Letters*, 46(21), 11850–11858.
- 545 Anantrasirichai, N., Biggs, J., Albino, F., & Bull, D. (2019b). A deep learning ap-  
 546 proach to detecting volcano deformation from satellite imagery using synthetic  
 547 datasets. *Remote Sensing of Environment*, 230, 111179.
- 548 Anantrasirichai, N., Biggs, J., Albino, F., Hill, P., & Bull, D. (2018). Applica-  
 549 tion of machine learning to classification of volcanic deformation in routinely  
 550 generated insar data. *Journal of Geophysical Research: Solid Earth*, 123(8),  
 551 6592–6606.
- 552 Banks, D., Davies, C., & Davies, W. (1995). The Chalk as a karstic aquifer: evi-  
 553 dence from a tracer test at Stanford Dingley, Berkshire, UK. *Quarterly Journal  
 554 of Engineering Geology and Hydrogeology*, 28(Supplement 1), S31–S38.
- 555 Bell, J. W., Amelung, F., Ferretti, A., Bianchi, M., & Novali, F. (2008). Perma-  
 556 nent scatterer insar reveals seasonal and long-term aquifer-system response  
 557 to groundwater pumping and artificial recharge. *Water Resources Research*,  
 558 44(2).
- 559 Blinowska, K. J., & Malinowski, M. (1991). Non-linear and linear forecasting of the  
 560 eeg time series. *Biological cybernetics*, 66(2), 159–165.
- 561 Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series  
 562 analysis: forecasting and control*. John Wiley & Sons.
- 563 Brockwell, P. J., & Davis, R. A. (2016). *Introduction to time series and forecasting*.  
 564 Springer.
- 565 Brown, R. G. (2004). *Smoothing, forecasting and prediction of discrete time series*.  
 566 Courier Corporation.
- 567 Burke, H., Hough, E., Morgan, D., Hughes, L., & Lawrence, D. (2015). Approaches  
 568 to inform redevelopment of brownfield sites: An example from the Leeds area  
 569 of the West Yorkshire coalfield, UK. *Land Use Policy*, 47, 21–331.
- 570 Carla, T., Intrieri, E., Di Traglia, F., & Casagli, N. (2016). A statistical-based  
 571 approach for determining the intensity of unrest phases at stromboli volcano  
 572 (southern italy) using one-step-ahead forecasts of displacement time series.  
 573 *Natural Hazards*, 84(1), 669–683.
- 574 Chacón, H. D., Kesici, E., & Najafirad, P. (2020). Improving financial time se-  
 575 ries prediction accuracy using ensemble empirical mode decomposition and  
 576 recurrent neural networks. *IEEE Access*, 8, 117133–117145.
- 577 Chambers, J., Weller, A., Gunn, D., Kuras, O., Wilkinson, P., Meldrum, P., ... oth-  
 578 ers (2008). Geophysical anatomy of the Hollin Hill Landslide, North Yorkshire,  
 579 UK. In *Near surface 2008-14th eage european meeting of environmental and  
 engineering geophysics* (pp. cp–64).
- 580 Chen, J., & Boccelli, D. L. (2018). Forecasting hourly water demands with seasonal

- 582 autoregressive models for real-time application. *Water Resources Research*,  
 583 54(2), 879–894.
- 584 Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk,  
 585 H., & Bengio, Y. (2014). Learning phrase representations using rnn encoder-  
 586 decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*.
- 587 Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. (1990). Stl: a  
 588 seasonal-trend decomposition. *Journal of official statistics*, 6(1), 3–73.
- 589 Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatter-  
 590 plots. *Journal of the American statistical association*, 74(368), 829–836.
- 591 Colesanti, C., Ferretti, A., Novali, F., Prati, C., & Rocca, F. (2003). Sar monitoring  
 592 of progressive and seasonal ground deformation using the permanent scatterers  
 593 technique. *IEEE Transactions on Geoscience and Remote Sensing*, 41(7),  
 594 1685–1701.
- 595 Crilly, M. (2001). Analysis of a database of subsidence damage. *Structural survey*,  
 596 19(1), 7–15.
- 597 Gaddes, M., Hooper, A., & Bagnardi, M. (2019). Using machine learning to auto-  
 598 matically detect volcanic unrest in a time series of interferograms. *Journal of*  
 599 *Geophysical Research: Solid Earth*.
- 600 Gaddes, M., Hooper, A., Bagnardi, M., Inman, H., & Albino, F. (2018). Blind signal  
 601 separation methods for insar: The potential to automatically detect and mon-  
 602 itor signals of volcanic deformation. *Journal of Geophysical Research: Solid*  
 603 *Earth*, 123(11), 10–226.
- 604 Greff, K., Srivastava, R. K., Koutnka, J., Steunebrink, B. R., & Schmidhuber, J.  
 605 (2017, Oct). Lstm: A search space odyssey. *IEEE Transactions on Neural*  
 606 *Networks and Learning Systems*, 28(10), 2222–2232.
- 607 Hamilton, J. D. (1994). *Time series analysis* (Vol. 2). Princeton New Jersey.
- 608 Hartmann, D. L., Michelsen, M. L., & Klein, S. A. (1992). Seasonal variations  
 609 of tropical intraseasonal oscillations: A 20–25-day oscillation in the western  
 610 pacific. *Journal of the atmospheric sciences*, 49(14), 1277–1289.
- 611 Heleno, S. I., Frischknecht, C., dOreye, N., Lima, J., Faria, B., Wall, R., & Kervyn,  
 612 F. (2010). Seasonal tropospheric influence on sar interferograms near the  
 613 itcz—the case of fogo volcano and mount cameroon. *Journal of African Earth*  
 614 *Sciences*, 58(5), 833–856.
- 615 Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Compu-*  
 616 *tation*, 9(8), 1735–1780.
- 617 Hylleberg, S. (1995). Tests for seasonal unit roots general to specific or specific to  
 618 general? *Journal of Econometrics*, 69(1), 5–25.
- 619 Hyndman, R. J., Khandakar, Y., et al. (2007). *Automatic time series for forecasting:*  
 620 *the forecast package for r* (No. 6/07).
- 621 Kingma, D.-P., & Ba, J. (2015). Adam: A Method for Stochastic Optimization. *In-*  
 622 *ternational Conference on Learning Representations (ICLR)*.
- 623 Lake, R., Northmore, K., Dean, M., & Tragheim, D. (1992). Leeds: a geological  
 624 background for planning and development: 1: 10000 sheets SE23NW, NE, SE  
 625 and SE33NW, NE, SW, SE: parts of 1: 50000 geological sheets 69 (Bradford),  
 626 70 (Leeds), 77 (Huddersfield) and 78 (Wakefield).
- 627 Lamont-Black, J., Younger, P. L., Forth, R. A., Cooper, A. H., & Bonniface, J. P.  
 628 (2002). A decision-logic framework for investigating subsidence problems po-  
 629 tentially attributable to gypsum karstification. *Engineering geology*, 65(2-3),  
 630 205–215.
- 631 large-scale insar monitoring of permafrost freeze-thaw cycles on the tibetan plateau.  
 632 (n.d.).
- 633 Lazecky, M., Hlavacova, I., Bakon, M., Sousa, J. J., Perissin, D., & Patricio, G.  
 634 (2016). Bridge displacements monitoring using space-borne x-band sar inter-  
 635 ferometry. *IEEE Journal of Selected Topics in Applied Earth Observations and*  
 636 *Remote Sensing*, 10(1), 205–210.

- 637 Lütkepohl, H., Krätzig, M., & Phillips, P. C. (2004). *Applied time series econometrics*. Cambridge university press.
- 638
- 639 Mazzanti, P., Rocca, A., Bozzano, F., Cossu, R., & Floris, M. (2011). Landslides
- 640 forecasting analysis by displacement time series derived from satellite insar
- 641 data: preliminary results. *Small*, 5000, 50–000.
- 642
- 643 McCay, A. T., Valyakakis, M., & Younger, P. L. (2018). A meta-analysis of coal
- 644 mining induced subsidence data and implications for their use in the carbon
- 645 industry. *International Journal of Coal Geology*, 192, 91–101.
- 646
- 647 Milillo, P., Giardina, G., DeJong, M., Perissin, D., & Milillo, G. (2018). Multi-
- 648 temporal insar structural damage assessment: The london crossrail case study.
- 649 *Remote Sensing*, 10(2), 287.
- 650
- 651 Rebane, J., Karlsson, I., Denic, S., & Papapetrou, P. (2018). Seq2Seq RNNs and
- 652 ARIMA models for cryptocurrency prediction: A comparative study. *SIGKDD*
- 653 *Fintech*, 18.
- 654
- 655 Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986, October). Learning repre-
- 656 sentations by back-propagating errors. *Nature*, 323, 533–.
- 657
- 658 Selvakumaran, S., Plank, S., Geiß, C., Rossi, C., & Middleton, C. (2018). Remote
- 659 monitoring to predict bridge scour failure using interferometric synthetic aper-
- 660 ture radar (insar) stacking techniques. *International journal of applied earth*
- 661 *observation and geoinformation*, 73, 463–470.
- 662
- 663 Sirignano, J., & Cont, R. (2019). Universal features of price formation in financial
- 664 markets: perspectives from deep learning. *Quantitative Finance*, 19(9), 1449–
- 665 1459.
- 666
- 667 Spaans, K., & Hooper, A. (2016). Insar processing for volcano monitoring and other
- 668 near-real time applications. *Journal of Geophysical Research: Solid Earth*,
- 669 121(4), 2947–2960.
- 670
- 671 Sutskever, I., Vinyals, O., & Le, Q. (2014). Sequence to sequence learning with neu-
- 672 ral networks. *Advances in NIPS*.
- 673
- 674 Taylor, S. J. (2008). *Modelling financial time series*. world scientific.
- 675
- 676 Torres, D. G., & Qiu, H. (2018). *Applying recurrent neural networks for multivariate*
- 677 *time series forecasting of volatile financial data*. Stockholm: KTH Royal Insti-
- 678 tute of Technology.
- 679
- 680 Valade, S., Ley, A., Massimetti, F., DHondt, O., Laiolo, M., Coppola, D., ... Wal-
- 681 ter, T. R. (2019). Towards global volcano monitoring using multisensor
- 682 sentinel missions and artificial intelligence: The mounts monitoring system.
- 683 *Remote Sensing*, 11(13), 1528.
- 684
- 685 Watson, K. M., Bock, Y., & Sandwell, D. T. (2002). Satellite interferometric ob-
- 686 servations of displacements associated with seasonal groundwater in the los
- 687 angeles basin. *Journal of Geophysical Research: Solid Earth*, 107(B4).
- 688
- 689 Weigend, A. S. (2018). *Time series prediction: forecasting the future and under-*
- 690 *standing the past*. Routledge.
- 691
- 692 Zubaidi, S. L., Dooley, J., Alkhaddar, R. M., Abdellatif, M., Al-Bugharbee, H., &
- 693 Ortega-Martorell, S. (2018). A novel approach for predicting monthly wa-
- 694 ter demand by combining singular spectrum analysis with neural networks.
- 695 *Journal of hydrology*, 561, 136–145.