Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 53-1 (2020) 628–633

Investigation of Mars seismic noise using modelling techniques developed on Earth

Kanchan Aggarwal * **Siddhartha Mukhopadhyay** ** **Arun K Tangirala** ***

* *Indian Institute of Technology Madras, Tamil Nadu, India (e-mail: kanchanagg08@gmail.com).*
** *Seismology Division, Bhabha Atomic Reseach Centre, Mumbai, India (e-mail: smukho64@gmail.com)*
*** *Indian Institute of Technology Madras, Tamil Nadu, India (e-mail: arunkt@iitm.ac.in).*

Abstract: Systems and signals theory has aided in understanding various kinds of processes such as meteorology, econometrics, engineering, etc. In this work, we explore the applicability of these theories to analyze phenomenon from other terrestrial bodies, specifically the seismic phenomenon on Mars, using data-driven approaches. It is expected that this analysis will provide insights into the formation and interior of the red planet. It will also allow us to draw a preliminary comparison of the seismic phenomenon between the home and red planet. The specific objectives of this work are (i) exploratory analysis of the statistical characteristics and (ii) develop a time-series model using a systematic procedure that was initially developed to analyze Earth data. It involves a rigorous investigation of the specific statistical properties, namely, stationarity, linearity, and Gaussianity, followed by the development of a suitable time-series model that is commensurate with these properties. Our analysis reveals that Mars noise exhibits specific types of non-stationarities, namely, trend and *heteroskedasticity*. In addition, noise also exhibits linearity and follows a Gaussian distribution. In line with these features, we develop a *component model* that comprises of a third-order polynomial trend and the Autoregressive Integrated Moving Average - Generalized AR Conditional Heteroskedasticity (ARIMA-GARCH) model. The findings reveal that the specific properties share a striking similarity with home planet data, while the trend appears to be a unique feature. These discoveries, we believe, will pave the way in understanding the red planet. The studies, including model development, are carried out on datasets recorded by the seismometer deployed on Mars during NASA's Insight mission.

Keywords: Mars seismic noise, heteroskedasticity, trends, time-series models

1. INTRODUCTION

System and signal theory has a universal impact in understanding different kind of processes from various spheres on Earth. Data-driven analysis of seismic data using these theories is essential for various reasons. Under-the-ground data not only helps in a better understanding of geology of a planet but also the physical properties of local and sub-shallow surfaces. Seismic data proves to be the most powerful tool to study the internal structure of planets. With this philosophy, several missions have been carried out to deploy seismometers on other terrestrial bodies. However, only two missions, Viking mission (1970s) and Interior Exploration using Seismic Investigations, Geodesy, and Heat Transport (InSight) mission (2018), were able to deploy seismometers successfully. Viking mission installed two seismometers on the red planet. Failed deployment of one of the seismometer and hampered recordings by another one during the Viking mission fail to give conclusive results. Furthermore, the second seismometer did not record any seismic event after 19 months of nearly continuous operation (Anderson et al. (1977)). It is because the seismometer was picking up vibrations from various operations of lander and Mars wind. These unexpected issues hampered the necessary geological data and the mission to uncover the crucial facts about the interior structure and composition of Mars.

Recently, Interior Exploration using Seismic Investigations, Geodesy and Heat Transport (InSight) discovery mission to Mars by NASA has deployed Seismic Experiment for Interior Structure (SEIS) seismometer onto the Mars surface using advanced robotic arm (Lognonné et al. (2019)). The InSight spacecraft landed on Mars on November 26, 2018, and since then, raw and uncalibrated seismic data is available online for research. Data collected for the first three months do not contain any marsquake or meteorite impacts. Analyzing this event-free data is necessary because, in many applications, background noise is assumed to exhibit certain properties.

Data collected from the Viking mission has been the interest of researchers for the past four decades. Several models of Mars has been proposed to study the interior structure. Okal and Anderson (1978) used the Preliminary

Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2020.06.105

Reference Earth Model (PREM) proposed by Dziewonski and Anderson (1981) as the reference model under the pressure conditions inside Mars to study the interior. Several other geochemical models were proposed (Zharkov and Trubitsyn (1978); Sanloup et al. (1999); Lognonné and Mosser (1993)) to have a better understanding of Mars structure. Apart from the theoretical models, there have been attempts to estimate the compositions of martian mantle (Plesa et al. (2015); Zheng et al. (2015); Khan and Connolly (2008)). However, these geochemical models generally fail to explain the geophysical data.

Despite the abundant literature to understand the interior structure of Mars, there have been minimal efforts to study the seismic noise. A better understanding of noise is very crucial in order to study the composition of the planet. Anderson et al. (1977) were among the first few researchers to focus on the Mars seismic noise. Analysis of Mars noise (Anderson et al. (1977); Lognonné and Mosser (1993)) shows that the seismic noise on Mars is competitively low and the primary source is the wind and the noise levels can be improved drastically by removing seismometers from the lander. Phillips and Grimm (1991); Anderson (1989) highlighted that the seismic properties are closely related to the interior structure of the planet. Therefore it is essential to characterize the noise systematically.

The main objective of this work is two-fold, (i) to systematically characterize the Mars seismic noise and (ii) develop a suitable time-series component model based on the statistical characterization of noise. Four prominent properties of noise are considered, namely, integrating effects or trends, heteroskedasticity, linearity, and Gaussianity. The first two properties correspond to specific types of non-stationarities, while the latter two are associated with the model structure and the distribution of driving force. The analysis also allows us to conduct a comparative study of the features exhibited by noise on Mars and Earth. The main focus of this work is to analyze how these properties vary for Mars and Earth data.

The primary findings of this work are that seismic noise from both the planets exhibits certain same features, while some of the properties are exclusive to Mars. Seismic noise from Mars also exhibit heteroskedasticity, but unlike the presence of integrating effects in Earth data, Mars noise exhibits a polynomial trend. Moreover, data from Mars tested positive for linearity and is found to be driven by Gaussian white noise (like Earth data), thereby calling for the widely used linear ARMA class of time-series models. ARMA models can model only the linear correlation and fail to model the variance changing nature of data. This heteroskedastic feature of noise is modeled using the GARCH model. Therefore, the overall model for the Mars seismic noise is a *component model* that consists of different types of models for different components of the data, (i) a third-order polynomial, (ii) ARMA model, and (iii) GARCH model.

The rest of the article is as follows. Section 2 reviews the theoretical definitions of necessary statistical properties and developing ARIMA-GARCH models. In Section 3, we describe the systematic method for characterizing the seismic noise followed by developing a suitable time-series model. Analysis of three data sets collected from the seismometer deployed at Elysium Planitia (ELYSE) site on Mars is illustrated using the proposed method is presented in Section 4. This Section also compares the seismic noise features on the home and red planet. The paper ends with a few concluding remarks in Section 5.

2. ESSENTIALS

This section lays down the theoretical foundations for the essential aspects that are used in the development of this work.

2.1 Definitions

For a random process $y[k]$,

1. *Wide-sense stationary process:* $y[k]$ is said to be weakly stationary (Tangirala (2014)) if it satisfies the following:
   (i) Mean of $y[k]$ is invariant with time.
   (ii) $y[k]$ has finite variance.
   (iii) Covariance between any two pair of observations of $y[k]$ is only a function of time distance, known as lag (l), and not the time.
2. *Heteroskedastic process:* $y[k]$ is said to be heteroskedastic if the second-order properties of $y[k]$, variance and spectral density, varies with time.
3. *Linear random process:* $y[k]$ is said to be linear if it can be explained as the linear combinations of driving force $e[k]$ (Shumway and Stoffer (1982)), i.e,
$$y[k] = \sum_{i=-\infty}^{\infty} c_i e[k-i], \quad \sum_{i=-\infty}^{\infty} |c_i| < \infty \tag{1}$$
   where, $e[k] \sim \text{iid}(0, \sigma_e^2)$.
4. *Gaussian random process:* $y[k]$ is said to be a Gaussian process if $\mathbf{Y} = \{y_1[k], y_2[k] \ldots y_n[k]\}$ for every collection of time n and every positive integer k have a non-singular multivariate normal distribution (Shumway and Stoffer (1982)).

2.2 Time-series models

The purpose of any time-series model is to capture the underlying characteristics of the process. In this work, we have used two different class of models, (i) ARIMA Brockwell et al. (2002) and (ii) GARCH (Engle (1982); Bollerslev (1986)) models. The former class of models are suitable for modelling the linear correlation in data while the later class models the time-varying variance in mean stationary data. Furthermore, ARIMA models are build on the data while GARCH are build on the residuals obtained from optimally estimated ARIMA model. Despite the difference in type of data used to model these classes of time-series models, the guidelines for developing a suitable model is same for both the classes. The generic procedure to develop a ARIMA model (Box et al. (2015)) is as follows:

(a) *Determine and model the non-stationarities in data.* The first step in the model development is to model the non-stationarities in data. There exists a variety of non-stationarities such as integrating effect, deterministic trend, periodicity, seasonality, etc. Each type of non-stationarity requires a specific kind of handling. Before estimating the model, data under consideration should be stationary.

(b) *Estimate the model of suitable order.* Stationary time-series can be modeled using $(p + m)$ parameters of the ARIMA model, values of p and m can be identified from auto-correlation function (ACF) and partial autocorrelation function (PACF) of data, which are estimated from the data using estimation algorithms such as least square, maximum likelihood method, etc.

(c) *Model assessment and validation.* The quality of the estimated model is assessed by the statistical analysis of residuals and the error analysis of estimates. Residual analysis ensures the whiteness of residuals, which is claimed if the ACF of residuals is insignificant at all lags, while the error analysis guarantees that the standard errors of the estimates are small as compared to the estimated values.

If the model does not meet the assessment test as mentioned above, then either the model structure or the order is refined, and the model is re-estimated until the estimated model passes the quality assessment step.

A linear stationary random process can be modeled using ARIMA(p, d, m), where an integrating type non-stationarity is modeled by differencing operation. For a time-series $y[k]$, $0 \leq k \leq N_x - 1$, mathematical formulation of ARIMA(p, d, m) is given by:

$$\Big(1 - \sum_{i=1}^{p} \phi_i q^{-i}\Big)(1 - q^{-1})^d y[k] = \Big(1 + \sum_{j=1}^{m} \theta_j q^{-j}\Big) e[k] \quad (2)$$

where, p and m are the orders of AR and MA, d represents the degree of differencing needed to model the integrating effect, q^{-1} is the back-shift operator and $e[k] \sim \text{GWN}(0, \sigma_e^2)$. As mentioned earlier, GARCH models are developed on the residual series $x[k]$ obtained from ARIMA model. Development of GARCH(P, Q) model follows same steps where the values of P and Q are identified from the ACF and PACF of squared residuals. A generalized-ARCH (GARCH) model of order (P, Q) is defined as:

$$\begin{aligned} x[k] &= \sigma_k \epsilon[k] \\ \sigma_k^2 &= c_0 + \sum_{i=1}^{P} b_i X_{k-i}^2 + \sum_{j=1}^{Q} a_j \sigma_{k-j}^2 \end{aligned} \quad (3)$$

where $P(\geq 1)$ and $Q(\geq 0)$ are the orders of ARCH and GARCH, $c_0 \geq 0$ represent the constant term, $b_i \geq 0$, $a_j \geq 0$ are the coefficients of estimated model. The driving force $\epsilon[k]$ is iid$(0, 1)$ and independent of $x_{k-l}, l \geq 1$ for all k. In the model assessment step, squared residuals are also tested for whiteness in addition to the GARCH residuals. One can observe from 3 that GARCH model is essentially an ARMA representation for σ_k^2 in terms of prediction errors and the variance of prediction error.

3. PROPOSED METHOD

In this section we present the procedure of a systematic methodology used to model the Earth's seismic noise. The outline of the methodology is shown in Figure. 1 As observed from Figure. 1, the procedure is implemented in two steps (i) on the seismic noise and (ii) on the residuals obtained from the time-series model.

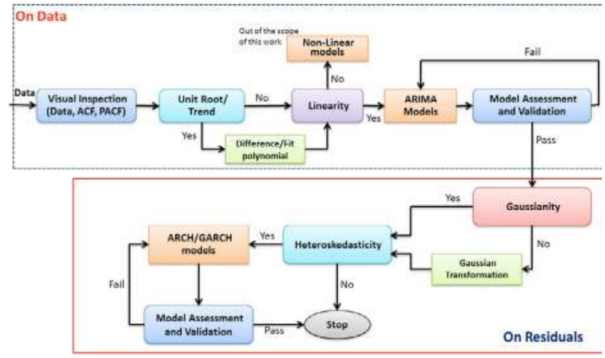

Fig. 1. Systematic Procedure to model Earth's seismic noise

3.1 Implementation on data

Given a dataset, test for the presence of integrating effect and linearity are implemented directly on data followed by the development of time-series models of suitable order.

Stationarity analysis The first step is to test the data for first-order stationarity conducted in two steps. First step being the visual inspection of various properties of data followed by the statistical test to support the inferences drawn from the visual inspection.

(i) **Visual Inspection:** Presence of trend or integrating type non-stationarity is assessed by inspecting the plots of data, autocorrelation function (ACF) and partial-correlation functions (PACF) of noise. Slowly decaying nature of ACF or the near-unity value of PACF at initial lags indicate the presence of integrating effect in data. Inferences drawn in this step are supported by well established statistical test.

(ii) **Statistical Tests:** Augmented Dickey Fuller (ADF) and Phillips-Perron (PP) tests are implemented to support the inferences of presence of integrating effect in seismic noise. In the presence of heteroskedasticity, test statistics of these test decreases resulting in the increased false rejection of null hypothesis. Because of the poor performance of these test for heteroskedastic data, there is a need to assess the presence of integrating effect in a different way. Another way to statistically confirm the presence of integrating effect is to fit an AR(1) model and analyze the estimated coefficient. A near-unity value of AR(1) coefficient confirms the presence of a pole at unit circle.

First-order non-stationarities are modeled before conducting the linearity test. Trend type non-stationarity is modeled either by de-trending the data or by fitting a polynomial of suitable order to the data while integrating type non-stationarity is modeled by differencing the data by suitable degree.

Linearity analysis First-order stationary data is tested for linearity using a surrogate-based hypothesis test. In this work, it is hypothesized that data is generated from a linear Gaussian process undergoing a static but nonlinear transformation and the time-series is quantified using correlation dimension (D_2) as test statistics. Null hypothesis, that the data is generated from linear Gaussian process, is rejected in the favour of alternate if D_2 of the original time-series differs from the ensemble of surrogates, i.e. correlation dimension of original data is compared with the empirical distribution of correlation dimension derived from surrogates. If the seismic noise exhibit nonlinear characteristic, then we restrict our model to the best linear approximated model at the cost of sub-optimal predictions.

Time-series model development and assessment After characterizing the seismic noise, next step is to develop a time-series model that is commensurate with the noise properties. Model development is carried out in two steps, (i) *model estimation* and (ii) *model assessment*. In this work, seismic noise is modelled using ARIMA(p, d, q) models of suitable order, where values of model order is selected based on ACF and PACF plots. The estimated model is selected for further analysis only if it passes the residual (whiteness) and error analysis (over-parameterization).

3.2 Implementation on residuals

Certain characteristics such as Gaussianity, presence of conditional heteroskedasticity, etc, assumes that the data is uncorrelated or weakly correlated. For this reason, statistical tests for such features cannot be performed on the original datasets. Whiteness of residuals claims zero correlation but fails to claim the independent nature of residuals. A process can be uncorrelated but dependent (referred to as ARCH effect) indicating the presence of heteroskedasticity in data. It is important to test the residuals for normality before testing for second-order stationarity because the linear time-series (ARIMA) models rest on the assumption of normal distribution. Moreover, statistical test for heteroskedasticity assumes data is generated from a linear Gaussian process.

Gaussianity analysis For colored data normality can be tested in two different ways:

(a) Testing directly on data: Using existing moments based test (Thode (2002)) or statistical test such as Shapiro-Wilk, Kolmogorov-Smirnov test, etc (Shapiro and Wilk (1965)).
(b) Testing on residuals after modeling the correlation structure in data.

We have adopted later approach since the former assumes that data is uncorrelated or weakly correlated which is not the case for seismic noise. The premise for conducting normality test on residuals is that since the white residual series is the forcing function for the linear model, its Gaussianity implies that the given seismic series is also jointly Gaussian. Normality of the residuals is claimed using the Shapiro-Wilk test which is the most powerful test among all the existing Normality test but is limited by the sample size (performance degrades drastically for sample size greater than 5000).

Second-order stationarity analysis Presence of second-order non-stationarity is also conducted in two steps, like first-order stationarity analysis. Heteroskedasticity in residuals is often referred to as the ARCH effect. Presence of ARCH effect can be guaranteed by analyzing the ACF of squared residuals. The correlation in squared residuals is the indication of presence of ARCH effect or heteroskedasticity in residuals. Inference drawn from visual inspection is supported by statistical test such as PSR and ARCH test.

Fig. 2. Location of Elysium Planitia (ELYSE) site

Heteroskedastic model development and assessment If the residuals tested positive for the presence of heteroskedasticity, then ARCH/GARCH models are used to model the heteroskedasticity in residuals, otherwise noise properties are reported with the estimated ARIMA model. Quality of the ARCH/GARCH model is assessed in the similar manner as for ARIMA model. In addition to whiteness of residuals, squared residuals are also analyzed to ensure unpredictability in the ARCH/GARCH residuals.

4. RESULTS AND DISCUSSIONS

The systematic noise modeling methodology is illustrated with the help of real-time Mars seismic noise downloaded from seismometers installed at Elysium Planitia (ELYSE) site on Mars. The data is freely available for research purpose on Incorporated Research Institutions of Seismology (IRIS). These seismometers installed during the InSight discovery mission to Mars by NASA. Details of the datasets are summarized in table 1, and the location of the seismometer is shown in Fig. 2. In this work, both the datasets are event-free, i.e., there is no seismic activity in the downloaded datasets. In order to develop time-series models, statistical characterization for both the datasets is carried out using well-established statistical tests for features such as first-order stationarity, linearity, heteroskedasticity, and Gaussianity. We also study the variation in these statistical properties with time.

Table 1. Specifications of data

Data	Network code	Station code	location code	Channel code	Date	start time	end time	sps
1.	XB	ELYSE	67	SHU	Jan 02,2019	00:00	01:00	20
2.	XB	ELYSE	67	SHU	Jan 09,2019	22:50	23:50	20
3.	XB	ELYSE	67	SHU	Jan 09,2019	11:50	12:50	20
4.	IU	ANMO	00	BHZ	Feb 27,2010	17:30	18:30	20

Application to dataset 1
Downloaded dataset 1 is shown in Fig. 3 along with the ACF and PACF of data. Visual inspection of data indicates the presence of trend-type non-stationarity while the slowly decaying ACF also indicates the presence of integrating effect. These qualitative inferences are tested statistically using ADF, PP, and KPSS tests. As observed from table 2, ADF and PP tests fail to support the claim of the presence of integrating effect in data while the KPSS test confirms the presence of a trend. Trend-type non-stationarity in data is modeled using a third-order polynomial. A red color line represents the estimated trend

in Fig. 3(a). First-order stationary data, along with its ACF and PACF, is shown in Fig. 4. ACF and PACF of processed data fail to indicate the presence of first-order non-stationarity in data. Statistical tests (table 2) also verify the claims drawn by visual inspection for processed data. First-order stationary data is tested for linearity using the surrogate-based approach. Table 2 shows that the test fails to reject the null hypothesis of linearity.

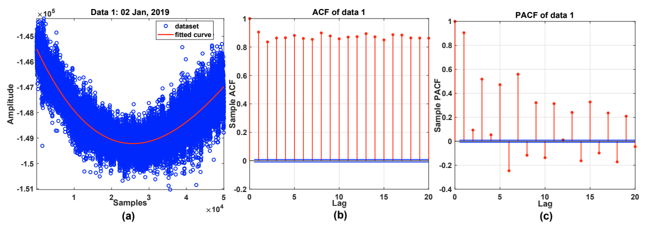

Fig. 3. Mars seismic (a) noise with third-order polynomial trend, (b) ACF, and (c) PACF of data

After ensuring the linearity and first-order stationarity of data, the next step is to develop an ARIMA model of suitable order. We use the Akaike Information Criterion (AIC) as the model selection criterion and estimate a range of models for varying AR and MA orders. Goodness of the estimated model is assessed by residual analysis (residuals exhibit no correlation, correlated residuals is the indication of under-fitting) and significance of estimated parameters (insignificant parameters indicate over-fitting). Our analysis reveals that ARIMA$(6, 0, 10)$ is a more suitable model for the Mars seismic data. Estimated model for dataset 1 ($y[k]$) is given by 4.

$$Dy[k] = P_k + Ne[k] \tag{4}$$

$$\begin{aligned}
P_k =& -110.3k^3 + 992.3k^2 - 95.54k - 1.492 \\
D =& \ 1 - \underset{(\pm.02)}{0.63}\, q^{-1} + \underset{(\pm.01)}{0.36}\, q^{-2} - \underset{(\pm.01)}{0.83}\, q^{-3} + \underset{(\pm.01)}{0.04}\, q^{-4} - \underset{(\pm.01)}{0.58}\, q^{-5} \\
& + \underset{(\pm.02)}{0.64}\, q^{-6} \\
N =& \ 1 + \underset{(\pm.02)}{0.93}\, q^{-1} - \underset{(\pm.01)}{0.72}\, q^{-2} - \underset{(\pm.02)}{0.68}\, q^{-3} - \underset{(\pm.01)}{0.84}\, q^{-4} - \underset{(\pm.03)}{0.87}\, q^{-5} \\
& + \underset{(\pm.04)}{0.49}\, q^{-6} + \underset{(\pm.02)}{0.76}\, q^{-7} + \underset{(\pm.01)}{0.33}\, q^{-8} - \underset{(\pm.01)}{0.05}\, q^{-9} - \underset{(\pm.01)}{0.15}\, q^{-10}
\end{aligned}$$

where, $\sigma_e^2 = 326$. The ACF of residuals (Fig. 5(a)) obtained from ARMA(6, 10) model reflects the white noise characteristics which is also indicated by the Box-Ljung test (Fig. 5(b)).

ARMA residuals are tested for Gaussianity and heteroskedasticity using SW and PSR or ARCH test, respectively. As observed from table 2, SW test fails to reject the null hypothesis that the residuals are generated from a Gaussian process. The test for heteroskedasticity

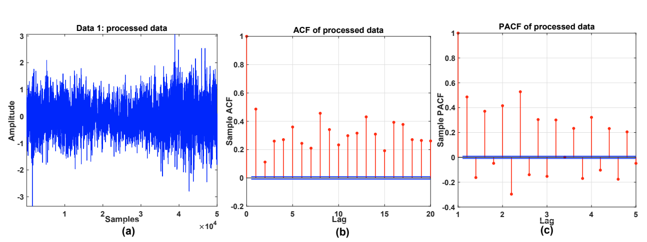

Fig. 4. (a) Trend-stationary seismic data, (b) ACF and (c) PACF of processed data

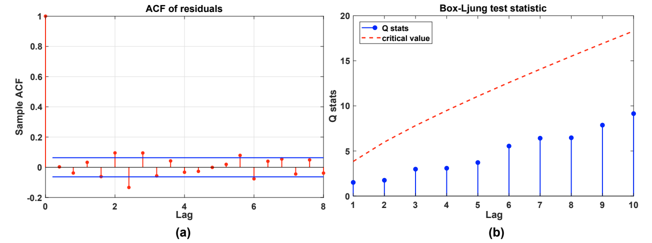

Fig. 5. (a) Residual ACF and (b) Box-Ljung statistics for ARMA(6,10) model

Table 2. Summary of statistical test for Mars seismic noise

Test	H0	p-value	Observed value	Lower critical value	Upper critical value	Nature of test	Outcome
Test on data							
ADF	Unit root is present	0.001	-131.49	-1.94	-	LT	RNH no unit root
PP	Integrating effects	0.001	-131.49	-1.94	-	LT	RNH no integrating effect
KPSS	Trend stationary	0.53	0.32	-	0.14	RT	Cannot RNH only trend
PSR **T** **I+R** **T+I+R**	Series is homoskedastic	0 0 0	3692.07 4843.09 8535.17	-	23.68 4009.73 4023.99	RT	RNH Heteroskedastic
Linearity	Series is linear	-	4.91	4.81	5.02	TT	Cannot RNH linear
Test on processed data							
ADF	Unit root is present	0.001	-131.49	-1.94	-	LT	RNH no unit root
PP	Integrating effects	0.001	-131.49	-1.94	-	LT	RNH no integrating effect
KPSS	Trend stationary	0.53	0.32	-	0.14	RT	RNH no trend
PSR **T** **I+R** **T+I+R**	Series is homoskedastic	0 0 0	3684.23 4853.83 8538.1	-	23.68 4009.73 4023.99	RT	RNH Heteroskedastic
Test on Residuals							
Shapiro-Wilk	Normal distribution	0.51	0.982	-	-	-	Cannot RNH Normal distribution
PSR **T** **I+R** **T+I+R**	Series is homoskedastic	0 0 0	3677.19 4849.29 8526.48	-	23.68 4009.73 4023.99	RT	RNH Heteroskedastic
ARCH	Series is homoskedastic	0	1181.12	-	3.84	RT	RNH Heteroskedastic

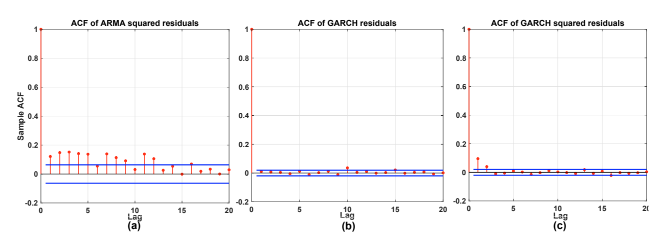

Fig. 6. (a) ARMA squared residuals ACF, (b) ACF of GARCH residuals and (c) GARCH squared residuals

follows test for normality. ACF of squared residuals (Fig. 6) indicates the presence of conditional heteroskedasticity in residuals which is statistically supported by both PSR and ARCH test (table 2). Conditional heteroskedasticity in residuals is modeled using GARCH models of suitable order. Assessment of a range of GARCH models suggests that GARCH(0,1) is the optimal model. Estimated GARCH$(0, 1)$ model for the ARMA$(6, 10)$ residuals is given by:

$$\begin{aligned}
e[k] &= \sigma_k z[k], \quad z[k] \sim N(0, 1) \\
\sigma_k^2 &= 24 + \underset{(\pm.03)}{0.1}\, e[k-1]^2
\end{aligned} \tag{5}$$

The GARCH residuals exhibit white noise characteristics. ACF of GARCH residuals and squared residuals are shown in Fig. 6. Our analysis shows that trend stationary seismic noise on MARS can be modeled as the combination of third-order polynomial trend and ARMA$(6, 10)$-GARCH$(0, 1)$ model.

Table 3. Summary of noise analysis for red and home planet

Data	Dataset 1	Dataset 2	Dataset 3	Dataset 4 (Earth data)
ADF &PP	No unit root	No unit root	No unit root	Unit root of order 2
KPSS	Trend stationary	Trend stationary	Trend stationary	No
Linearity	Yes	Yes	Yes	No Depends on location
Gaussianity	Yes	Yes	Yes	Yes
PSR & ARCH	Heteroskedastic	Heteroskedastic	Heteroskedastic	Heteroskedastic
Trend model	3 order polynomial	3 order polynomial	3 order polynomial	No
Model	ARMA(6,10)-GARCH(0,1)	ARMA(6,10)-GARCH(0,1)	ARMA(6,10)-GARCH(0,1)	ARIMA(5,2,3) - GARCH(1,1) order varies the location

Application to dataset 2 *and* 3

A similar procedure is applied to dataset 2 and 3. It is observed that the noise properties for these datasets are similar as for the dataset 1 (due to page limitation we have only summarized the results without the details in table 3). Thus, the seismic noise can be explained using the same model structure (a polynomial trend with ARMA$(6, 10)$-GARCH$(0, 1)$ model) as for dataset 1. However, the estimated parameters are different for all the datasets. Variation in estimated parameters is expected as the primary source of noise are likely to vary in 24-hour period.

We had carried out a similar analysis for Earth's seismic noise. Interestingly, our analysis from Earth and Mars data reveals that noise from both the planets exhibit certain similar features such as linearity, heteroskedasticity, and Gaussianity while a few of the features observed in Earth's data are missing in Mars data such as the presence of integrating effects. Surprisingly, all the datasets from Mars exhibit polynomial trend, which is a unique characteristic of Mars data and is absent in Earth's data (we have analyzed approximately 400 datasets from our planet). Based on the analysis carried out in this work, we speculate that both the planets have similar interior structure (crust, mantle, and core) as the datasets exhibit similar statistical properties, but the types of rocks that form the planet is perhaps different. It can be the rocks or the depth of various layers that are resulting in the different properties for these planets. It can also be speculated that the trend or integrating type non-stationarities are due to the instrumentation and has nothing to do with the surface of the planets.

5. CONCLUSION

The primary purpose of this work was to establish a link between the interior structure of Earth and Mars based on the data-driven analysis of seismic noise. Therefore, we investigated the seismic data from the red planet (Mars) using the systematic methodology proposed to analyze seismic noise from Earth. The significant discoveries of this work are two-fold, (i) Mars noise exhibits exciting features such as polynomial trend, linearity, heteroskedasticity, and Gaussianity. The noise can be modeled using a trend with ARMA-GARCH class of models (ii) noise from both the red planet and our home planet exhibits certain similar features while some of the characteristics are unique to both the planets. These features reflect the underlying similarity in the structure of both the planets. However, how these properties can be associated with the study of the interior structure of planets is still a question which needs the understanding of the underlying mechanism of seismic noise generation.

REFERENCES

Anderson, D.L. (1989). *Theory of the Earth*. Blackwell scientific publications.

Anderson, D.L., Miller, W., Latham, G., Nakamura, Y., Toksöz, M., Dainty, A., Duennebier, F., Lazarewicz, A.R., Kovach, R., and Knight, T. (1977). Seismology on mars. *Journal of Geophysical Research*, 82(28), 4524–4546.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3), 307–327.

Box, G.E., Jenkins, G.M., Reinsel, G.C., and Ljung, G.M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.

Brockwell, P.J., Davis, R.A., and Calder, M.V. (2002). *Introduction to time series and forecasting*, volume 2. Springer.

Dziewonski, A.M. and Anderson, D.L. (1981). Preliminary reference earth model. *Physics of the earth and planetary interiors*, 25(4), 297–356.

Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987–1007.

Khan, A. and Connolly, J. (2008). Constraining the composition and thermal state of mars from inversion of geophysical data. *Journal of Geophysical Research: Planets*, 113(E7).

Lognonné, P., Banerdt, W.B., Giardini, D., Pike, W., Christensen, U., Laudet, P., De Raucourt, S., Zweifel, P., Calcutt, S., Bierwirth, M., et al. (2019). Seis: Insight's seismic experiment for internal structure of mars. *Space Science Reviews*, 215(1), 12.

Lognonné, P. and Mosser, B. (1993). Planetary seismology. *Surveys in Geophysics*, 14(3), 239–302.

Okal, E.A. and Anderson, D.L. (1978). Theoretical models for mars and their seismic properties. *Icarus*, 33(3), 514–528.

Phillips, R.J. and Grimm, R.E. (1991). Martian seismicity. In *Lunar and Planetary Science Conference*, volume 22.

Plesa, A.C., Tosi, N., Grott, M., and Breuer, D. (2015). Thermal evolution and urey ratio of mars. *Journal of Geophysical Research: Planets*, 120(5), 995–1010.

Sanloup, C., Jambon, A., and Gillet, P. (1999). A simple chondritic model of mars. *Physics of the Earth and Planetary Interiors*, 112(1-2), 43–54.

Shapiro, S.S. and Wilk, M.B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4), 591–611.

Shumway, R.H. and Stoffer, D.S. (1982). An approach to time series smoothing and forecasting using the em algorithm. *Journal of time series analysis*, 3(4), 253–264.

Tangirala, A.K. (2014). *Principles of system identification: theory and practice*. Crc Press.

Thode, H.C. (2002). *Testing for normality*. CRC press.

Zharkov, V. and Trubitsyn, V. (1978). Physics of planetary interiors, wb hubbard, ed.

Zheng, Y., Nimmo, F., and Lay, T. (2015). Seismological implications of a lithospheric low seismic velocity zone in mars. *Physics of the Earth and Planetary Interiors*, 240, 132–141.