Examining satellite images market stability using the Records theory: Evidence from French spatial data infrastructures

Chady Jabbour[1], Anis Hoayek[2], Pierre Maurel[3], Zaher Khraibani[4], and Latifa Ghalayini[5]

[1]CEE-M, Univ. Montpellier, CNRS, INRAE, Institut Agro, France / Faculty of Economic Science and Business Administration, Lebanese University, Lebanon
[2]UMR 6158 LIMOS, Institut Henri Fayol, Mines Saint-Etienne, Univ. Clermont Auvergne, CNRS, France
[3]TETIS, INRAE, AgroParisTech, CIRAD, CNRS, Univ. Montpellier, France
[4]Department of Statistics, Faculty of Sciences, Lebanese University, Lebanon
[5]Faculty of Economic Science and Business Administration, Lebanese University, Lebanon

---

Abstract: The spatial data infrastructures (SDIs) which constitute a direct link between spatial data users and the large Earth observation industry, have a leading role in establishing market opportunities in the space sector. The spatial information supplied through various forms of SDI platforms exhibits large increases in demand volatility. The users' demand is unpredictable and the market is vulnerable to high evolution shifts. We study the effect of extreme demands for a particular type of spatial information, the satellite images. Drawing on two French SDIs, GEOSUD and PEPS, we examine the shifts occurring on their platforms and assess the probability of witnessing a spike/drop in the short term of different satellite imagery schemes: the high resolution images through GEOSUD; the Landsat (U.S.), Sentinel (Europe) and SPOT (France) images through PEPS. We analyze the market stability through the two SDIs and evaluate the probability of future records by using the Records theory. The results show that the high resolution images demand through GEOSUD, for which the classical i.i.d. model fits the most, is stable. Moreover, the Yang-Nevzorov model fits to the Landsat data, due to more records concentrated beyond the first observations. The Landsat demand is the less stable out of the other three satellite images series, and the probability of having a record in the coming years is the highest. While the use of Records theory drops mathematical constraints, it offers an alternative solution to the non-applicability of the machine learning techniques and long-term memory models.

Keywords: Landsat, market stability, Records theory, satellite images, spatial information, spatial data infrastructure, spot, Yang-Nevzorov

1 Introduction

The Earth observation (EO) market is comprised of several multi-scale markets (local, national, international, etc.) largely developed over the last four decades [8]. Since the launch of Landsat in 1970s, various types of EO data have seen the light. Recently, the EO imagery landscape is being shared by two main providers: the Digital Globe (U.S.), considered today as the worldwide leader in EO data and services and Airbus Defence and Space (Europe) [55]. Despite the particularity of each provider's profile, several common tasks are shared, consisting mainly of delivering satellite imagery and operating direct receiving stations [20]. The majority of their revenues results from image sales and services delivery. Therefore, combining satellite imagery with other types of data, is one of the main objectives in order to create useful information for end-users [39].

In such context, the users are becoming primary key-drivers for this technology [13]. They contribute through their demand of raw data and services, to its development and growth [52]. The spatial data infrastructures (SDIs), which constitute the direct link between the spatial data users and the large EO industry, have a leading role in establishing market opportunities [19]. In fact, the demand via these SDIs, depends on several factors: the role of the users' critical mass [23], the services available to accompany the raw data [51], the level of openness of the SDI platforms, etc. [12]. It is thus challenging to get an assessment of the global EO market demand, and more particularly the satellite image market; several factors may be of influence. First, the increasing number of data providers/suppliers make it more complicated to have a precise view over the global market-chain [53]: there exist over seventy satellite-based observing programs operated by more than thirty different countries [8], with markets structured through complex interoperability between upstream and downstream layers [62]. Secondly, the redistribution of value between raw data and added-value services and products: data users are less inclined to pay for raw or pre-processed data, even at preferential rates, due to existing upstream pooling mechanisms put in place (e.g., the Copernicus European program) and the scarcity of public budgets [5,16]. This phenomenon affects the value chain and consequently, the demand side. Third, the increasing availability of free images have also a direct effect on the overall images price in the market. The Landsat case comes as is an example where the images are publicly available for free [68,69]. Finally, the images' quality. Up to now, the satellite images prices are mainly driven by their quality [27].

All these elements raise the competition levels between the satellite image markets and highlight the questioning about the users' demand. The satellite images, which are traded either via commercial operators or supplied through other forms of platforms (two-sided market platforms, public-private partnerships, etc.), exhibit large increases in demand volatility [34]. The satellite imagery demand is unpredictable and the market is vulnerable to high evolution shifts. While globally the EO market was worth $50 billion in 2017, it is expected to reach $80 billion by 2022. Additionally, the added-value services are expected to grow from U.S. $28 billion in 2017 to some U.S. $65 billion estimated by 2022 [24]. In such circumstances, and face to these facts, studying the satellite image market could take several forms.

The objective of this paper is to study the stability of different satellite image markets through two independent French SDIs and forecast the users' future demand using the Records theory. In fact, the Records theory has been applied in several fields. Form climate change, natural gas markets to finance and sports, this method has proved its relevance in

capturing the dynamics occurring in certain markets or through sets of data [31]. In our case, the study will be based first on the HR satellite images demand, through the GEO-SUD[1]/Theia[2] SDI, in France. Since 2011, the GEOSUD and Theia Land data center form a unique SDI, aiming to progressively build-up an ecosystem of innovation in the field of satellite imagery for Earth observation. On the other hand, and more generally, three satellite image markets through the PEPS[3] SDI platform will be also explored [Landsat (U.S.), Sentinel (Europe) and SPOT (France)]. Designed since 2015, the PEPS platform is managed by the French National Space Agency (CNES). It constitutes the direct link between the European Copernicus program and the French user needs at all levels, from local to national.

We will assess the probability of witnessing a spike/drop in the short term in these different markets. To the authors' knowledge, this work is the first application of the recent advances of the Records theory to the geospatial science field, and more particularly, the SDI satellite images. It comes in a continuity of the reflections deployed within the GIS community, to offer an understanding of the dynamics that could occur through the users' geospatial demand. Hence, leading to a better comprehension of the management of this data, the fluctuations that the SDI could face on its platform, and thus its market stability. Worthy to note that the choice of the three satellite image markets for which we had a privilege access through the PEPS platform statistical data, occupies a unique place in the EO pantheon. The Landsat is the longest running uninterrupted Earth observation program for acquisition of satellite imagery of Earth's land [41]. It is recognized as one of the largest U.S. Earth observation programs, with an estimated annual benefit value of nearly $1.8 billion [45]. Concerning the Sentinel satellite constellation, it is considered as the most important EO program developed by the European Space Agency (ESA) to support the European operational and policy needs of the Global Monitoring for Environmental Security (GMES) [4, 21, 64]. With its long-term continuity of measurements for more than twenty-year's time, its global frequent coverage and its broad variety of sensing methods [9], it is designed to provide routine observations for operational Copernicus services. Finally, the SPOT commercial high-resolution EO satellites managed by the CNES in France, has already taken more than 10 million high-quality images since its beginnings in 1986. It was designed to improve the knowledge and management of the Earth by exploring the Earth's resources and monitoring human activities and natural phenomena.

Although the stability of the satellite images markets cannot be dissociated from previously mentioned factors, the literature examining the technical issues related to the expansion of the EO markets [1, 55] has not addressed clearly how the users' demand can reveal part of the market risk [20, 22]. Studying the users' demand occurring via the SDI platforms still presents many complications [44, 57, 66, 71]. Usually, the time series of the satellite images demand are non-*iid*, and assumptions about the type of distribution are complex. Among the various extreme value theory researches, analyzing the high-frequency time series data or the evidence of causality between the supply and demand variables relies basically on stochastic econometric schemes [40]. These models contain a large number of parameters, a fact that poses estimation challenges and over-parameterization concerns [17, 28]. Such difficulties do not concern the Records theory which drops several mathematical constraints about the choice of the underlying distribution and the quality of

---

[1]GEOSUD web site: http://equipex-GEOSUD.fr

[2]Theia web site: http://www.theia-land.fr/en

[3]http://peps.cnes.fr

the residuals [31]. Additionally, it could be used in cases where the number of available observations is somehow small [2]; usually, the study of time in record models is accounted through particular random variables called 'record indicators' [6]. Moreover, this theory offers an alternative solution to the non-applicability of the machine learning techniques and long-term memory models, where a considerable amount of data is needed for a good performance [46, 67].

The contribution of this paper is in presenting a new approach for advancing the geospatial science. The novelty relies in offering additional processing elements in order to enrich the GIS market policy development, by bringing a new perspective on possible risks and facts that could be present while considering the geospatial markets dynamics. The cases of the GEOSUD and PEPS SDIs, for which the analysis was performed, serve as relevant illustration for the proposed approach. Through the framework presented, a grounded mathematical modeling tool is combined to to the usage of geospatial information, with a focus on the market forecasting analysis techniques. The remainder of the paper is organized as follows. We ïn̈A̧rst outline the methodological protocol used in this study (description of the SDIs characteristics, satellite-images schemes, data collection process). In a second time, we present the record models used in our work (classical and Yang-Nevzorov) followed by a statistical estimation of parameters and a goodness of fit test analysis. Next, the discussion brings those elements together. We test the models, check the reliability of the results, explain their significance and highlight their impacts. Finally, we conclude the article by considering future prospects in this field.

2   Methodology

2.1   SDI characteristics'—GEOSUD / PEPS

GEOSUD aimed since its launch at the development of the geospatial technology and data with a particular focus on the high resolution satellite images, contributing to its availability and easy access for its direct users. While facilitating the creation of integrated EO environment and supporting the geospatial communities in developing their ecosystem, it is continuously exploiting new opportunities through the steady development of its platform. It is targeting both the scientific community and the public actors based on a concept of open innovation. Through developing a set of functionalities for the accompaniment of its users, GEOSUD is lowering the barriers and reconsidering the value chain of the geospatial data, with all the economic and societal benefits that may emerge behind.

On the other hand, PEPS constitutes the missing link between the European spatial infrastructure and the French institutional users, scientists and industries. Managed by the CNES, it is conceived to offer innovative services and open access, mainly to the Sentinel satellite images and products. By redistributing geospatial data and enabling users to process them on servers close to the source, it is helping in achieving Copernicus goals in implementing and monitoring environmental and security policies. As a concrete example, the Theia scientific community is increasingly using PEPS via the CNES system to directly access to the whole Sentinel archive and more particularly Sentinel-2A data, in order to make atmospheric corrections (level 2A) and monthly syntheses (level 3A).

2.2   Satellite-images schemes

2.2.1   GEOSUD high resolution satellite imagery

In November 2018, more than 15,000 images, covering 55 million sq.km., have been downloaded by GEOSUD direct users since the beginning of the project. The HR images provided through the SDI web portal vary mainly between Rapid Eye and SPOT 1-5. Lately, GEOSUD extended its image acquisition capacity by settling a new Antenna in Montpellier (south France), with the aim of accomplishing by itself the SPOT 6/7 image production process. It is responding to the needs of the scientific community, through a pooling system developed specifically to better sharing the benefits of the open innovation processes. For instance, an absence of such a system (an investment of €11 million), would have cost €110 million to GEOSUD users to acquire images separately from independent spatial data providers at preferential rates [33].

2.2.2   PEPS satellite imagery: Landsat, Sentinel, and SPOT

While PEPS has been designed initially to offer an access to Sentinel satellites images, other satellite images are also present on the platform:

- First, the Landsat images. Landsat was the first EO system to employ a global image acquisition strategy, offering satellite imagery in a free and open manner [68]. The USGS Landsat data archive currently contains almost 6 million images [70], with an actual acquisition capacity with the Landsat 7 and 8 of approximately 1200 scenes per day.
- Second, the Sentinel images. Although the Sentinel images cannot be considered a replacement or an alternative for Landsat images, they offer a promising augmentation to the Landsat program [18]. From global warming to land use change and the atmosphere, the Sentinel family of satellites offers free of charge data to encourage its maximum use. It aims at achieving a global, continuous and wide range of Earth observation.
- Third, the SPOT images. The SPOT system consisting of five satellites launched since 1986, has produced more than 30 million images, ready to be used in studying environmental change over more than a quarter-century. Similarly to the U.S. government's experience in opening spatial data access (with its decision to no longer charge for access to the Landsat satellite imagery), the CNES has recently opened its SPOT Earth observation data archive through the PEPS platform [50].

2.3   Data collection

The data collection process via the two SDIs platforms (PEPS & GEOSUD), as well as its pre-analysis activity, were performed as follows:

- First, key-contact persons were identified in each of the SDIs (GEOSUD: platform operators, PEPS: download data managers), in order to respond to our various data requests.
- Second, meetings were performed with the identified persons in order to agree on data types, quality and formats. After several exchanges, the desired data were collected in its row format as excel sheets (due to difficulties in cleaning data at source
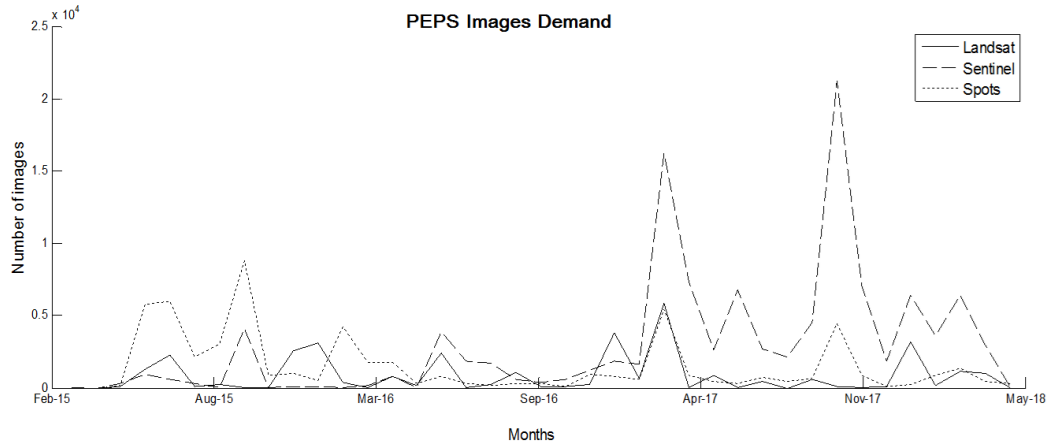
Figure 1: Monthly demand of the three PEPS satellite imagery schemes.

and lack of means and time). The collected data consisted of a compiled file, group-
ing various variables related to the downloaded images (including duplicated down-
loads, registered download failures, etc.).

- Third, in order to make the data usable, thorough analysis was performed with the
  help of data scientists and geomatics experts, due to the complexity encountered in
  terms of satellite databases sorting; and
- Finally, the new database was ready to be used and tested under the several theoreti-
  cal models proposed (see Section 2.4).

Users demand for the Landsat, Sentinel and SPOT images were collected between
February 2015 (the launching date of the platform) and May 2018. Figure 1 shows a time
series of the monthly satellite image demand of the three different markets via the PEPS
platform.

Early 2015, the demand for Sentinel was quite low as shown in Figure 1. A fact justified
by the first launch of the Sentinel satellites (Sentinel-1A in 2014, Sentinel-2A in 2015). How-
ever, in 2016, four Sentinel satellites were already in orbit, including Sentinel-1B in April
2016. As a direct consequence, a slight increase in demand could be noticed during the
period following April 2016. On the other hand, the two peaks in 2017 could be explained
by two phenomena:

- First, the launch of two additional satellites Sentinel-2B in March 2017 and Sentinel-5
  in October 2017, and its direct impact on the growing demand for images by the PEPS
  platform users.
- Second, it is important to specify that these two peaks precede two important peri-
  ods within the French academic year and research institutional sphere (universities,
  research institutes, research centers, etc.). It includes setting up internship projects for
  Master students (March–April) as well as thesis and research projects for Ph.D. stu-
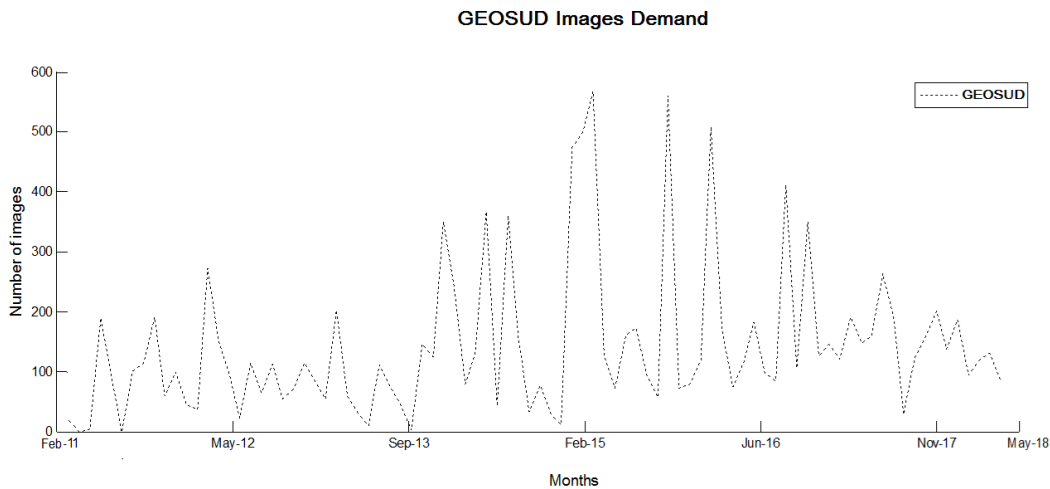
Figure 2: Monthly demand of the GEOSUD HR satellite imagery schemes.

dents and researchers (September–October). Both events are often accompanied by massive demands for images, and particularly when projects are based on geospatial information (management of natural resources, agriculture, land use planning, etc.).

Concerning the GEOSUD HR images, the data was recorded for a longer timeframe, going from February 2011 till May 2018, because of its availability since the beginning of the GEOSUD project. Table 1 summarizes the PEPS and GEOSUD data sets used in this study.

Satellite images program	Frequency	Number of observations	Web portal
Landsat	Monthly	39	PEPS
Sentinel	Monthly	39	PEPS
SPOT	Monthly	39	PEPS
GEOSUD HR images	Monthly	87	GEOSUD/Theia

Table 1: GEOSUD and PEPS data sets.

The monthly demand of the GEOSUD HR satellite images are illustrated in Figure 2. We will interpret these graphs statistically and economically by using the Records theory in the next section.

2.4   Records theory model—general context

The study of records initiated with Chandler [15]. The first results obtained, defined as "classical records model," presents the case where the underlying random variables (*rv*) are independent, identically distributed (*iid*). Nevzorov [48] and Arnold et al. [3], through their collaboration, brought significant development to the initial record literature. Later

on, recognizing the unavailability of these results to fit several sets of data, an effort was made in order to go beyond the context of classical records, where the observations are independent but not identically distributed. Interested readers can consult the works of Nevzorov [49] for further details.

We begin by introducing the general context, and then present some useful results in the classical and non-*iid* cases:

Let $\{X_t, t \geq 1\}$ be a sequence of *iid* random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. $F(\cdot)$ and $f(\cdot)$ are respectively the corresponding cumulative distribution and density functions. In a chronological series, an observation X_j is said to be an upper record if it is higher than all the previous observations, i.e. $X_j > \max\{X_1, \ldots, X_{j-1}\}$. In this paper, without loss of generality, we consider only the upper records, knowing that the lower records can be defined in a similar way (e.g., by multiplying the chronological series by "-1"). Following the increase of t, the value and the occurrence index of the nth record are respectively given by the following sequences $\{R_n, n \geq 1\}$ and $\{L_n, n \geq 1\}$, i.e. $R_n = X_{L_n}$. In such context, the available data is a sequence of the couples $\{(R_n, L_n), 1 \leq n \leq N_T\}$, where T is the present time and N_T is the number of records in the sequence $\{X_t, 1 \leq t \leq T\}$.

In fact, there exists a similarity between "records" and "processing of censored data" [31]. To highlight this similarity, we consider the sequence of record indicators $\{\delta_t, 1 \leq t \leq T\}$ defined by:

$$\delta_t = \begin{cases} 1 & \text{if } X_t \text{ is a record} \\ 0 & \text{otherwise} \end{cases}$$

Therefore, based on the record indicators, a mathematical expression of the number of records is given by:

$$N_T = \sum_{t=1}^{T} \delta_t. \tag{1}$$

2.4.1 Classical model

Several stochastic properties of the record series in the *iid* case are distribution-free, i.e. they fit to any underlying distribution of X_t. We present in this section, the most popular results in the *iid* context that will be used later in our work.

Nevzorov [49] showed that $\forall T \geq 1$, the record indicators $\delta_1, \ldots, \delta_T$ are independent and each δ_t is distributed according to a Bernoulli distribution of parameter $\frac{1}{t}$, i.e.

$$P_t \equiv \mathbb{P}[\delta_t = 1] = \frac{1}{t}. \tag{2}$$

Where, P_t is the probability that X_t is a record, defined as "record rate" at time t. Based on Equation (1) and on the previous result of Nevzorov [49], the expected value and the variance of the number of records are expressed by:

$$\mathbb{E}[N_T] = \sum_{t=1}^{T} \mathbb{E}[\delta_t] = \sum_{t=1}^{T} P_t = \sum_{t=1}^{T} \frac{1}{t}, \tag{3}$$

and

$$\mathbb{V}[N_T] = \sum_{t=1}^{T} \mathbb{V}[\delta_t] = \sum_{t=1}^{T} P_t(1 - P_t) = \sum_{t=1}^{T} \frac{1}{t} - \sum_{t=1}^{T} \frac{1}{t^2}. \tag{4}$$

Based on Equation (2), the record rate converges asymptotically to zero. Hence, in the *iid* case, when t increases, the records tend to become more distant in time. This implies in general that records are concentrated among the first observations. However, in several real data sets, this phenomenon is not always true for many reasons. For example, the technological progress can make the occurrence of records more frequent than expected by the iid model. This has led to the development of more comprehensive models that can provide better prediction. In the next section, we go beyond the *iid* case by considering the Yang-Nevzorov model [71,72].

2.4.2 Beyond *iid*—Yang-Nevzorov

The Yang-Nevzorov model is considered the most popular in the non-*iid* context due to several reasons:

1. For having the structure of a proportional risk model in survival analysis, which has shown its usefulness in modeling many datasets [31];
2. For generalizing the results of the Linear Drift Model (LDM) introduced by Ballerini and Resnick [6] which also represents a case beyond the *iid*.

In this model, the observations are assumed to be independent but not identically distributed. In order to adjust a Yang-Nevzorov or LDM model to a set of data, an additional step is needed. One approach consists of testing the null hypothesis:

$$H_0: \text{the data derive from a series of } \textit{iid rv}.$$

Such an adequation test could be based on the sequence of the record indicators $\delta_1, \ldots, \delta_T$ by using the theorem of Arnold and Villaseñor [3]. The theorem states that under the null hypothesis, $\mathcal{N}_T$, representing a standardized version of the number of records, converges to the standard normal distribution $N(0,1)$. $\mathcal{N}_T$ is given as follows:

$$\mathcal{N}_T = \frac{N_T - \log T}{\sqrt{\log T}}. \tag{5}$$

As the Yang-Nevzorov model implies an increase in the number of records, H_0 is rejected if

$$\mathcal{N}_T > z_{1-\alpha},$$

where $z_{1-\alpha}$ is the $(1-\alpha)^{th}$ quantile of $N(0,1)$ (α is the type one error usually set at 5%).

In the Yang-Nevzorov model introduced by Yang [72], a fixed integer ρ_t of *iid rv Y* of cumulative distribution function (CDF) $F(\cdot)$ is generated simultaneously at time t, from which $X_t = \max(y_1, y_2, \ldots, y_{\rho_t})$ is extracted. Thus, the sequence $\{X_t, t \geq 1\}$ of independent but not identically distributed random variables is considered, with the following CDF:

$$F_{X_t}(x) = F(x)^{\rho_t}, \rho_t > 0.$$

Concerning the parameter ρ_t we adapt the parametric form introduced originally by Yang [72] and developed later by Nevzorov [49]. Let ρ_t take real values,

$$\rho_t = \gamma^t, \text{with } \gamma > 1. \tag{6}$$

In this case ρ_t represents an exponential growth in the number of available *rv*. This parametric form has shown a great utility in various fields of application, especially in survival analysis models and in the detection of records in a context where the Yang-Nevzorov has been also used [38]. Hence, based on the *iid* property of the generated *rv*, and using the parameterization (6), the rate of record related to the sequence X_t, at time t, is:

$$P_t = \mathbb{P}[\delta_t = 1] = \frac{\rho_t}{\sum_{k=1}^{t} \rho_k}, t \geq 1,$$

$$= \frac{\gamma^t(\gamma - 1)}{\gamma(\gamma^t - 1)}. \tag{7}$$

Therefore, in a Yang-Nevzorov model, we can define an asymptotic rate of record by letting t go to infinity:

$$\lim_{t \to \infty} P_t \equiv P(\gamma) = \frac{\gamma - 1}{\gamma}. \tag{8}$$

Note that, the existence of an asymptotic rate of record is interpreted in the sense that the occurrence of a record is always possible even in the long run forecast. This case is generally related to a context where the studied *rv* represents high instability, i.e. a random variable showing high volatility levels. In addition, as in the *iid* classical case, Nevzorov [48] showed that the indicators are independent and follow a Bernoulli distribution of parameter P_t. This property is true under any underlying distribution (i.e., the distribution of Y). Therefore, this type of property is called distribution free. Hence, based on Equation (1) and on the previous distribution free property of Nevzorov [48], the expected value and the variance of the number of records, in the Yang-Nevzorov context, are expressed by:

$$\mathbb{E}[N_T] = \sum_{t=1}^{T} P_t \text{ and } \mathbb{V}[N_T] = \mathbb{E}[N_T] - \sum_{t=1}^{T} P_t{}^2. \tag{9}$$

2.4.3   Distribution free estimation of γ

As long as the Yang model parameter γ is unknown, the model cannot be used in practice; an estimation of γ is needed. Hoayek et al. [31] showed that the best estimator, in terms of bias and standard deviation, is obtained by applying the maximum likelihood principle. Using the independence property of the record indicators, the log likelihood function to be maximized is given as follows:

$$\ln L(\gamma) = N_T \ln\left(1 - \frac{1}{\gamma}\right) + (T - N_T) \ln\frac{1}{\gamma} - \ln\left(1 - \frac{1}{\gamma^T}\right) - \sum_{t=2}^{T} \delta_t \ln\left(1 - \frac{1}{\gamma^{t-1}}\right). \tag{10}$$

By maximizing Equation (10) with respect to γ, we obtain the distribution free maximum likelihood estimator denoted by $\widehat{\gamma}$. Moreover, Hoayek et al. [31] studied the asymptotic behavior of the estimator, which is necessary for constructing the estimatorâĂŹs asymptotic confidence intervals:

$$\frac{\widehat{\gamma} - \gamma}{\sqrt{I_T^{-1}(\gamma)}} \rightarrow N(0,1), \tag{11}$$

Where $I_T^{-1}(\gamma)$ represents the Fisher information.

2.4.4 Goodness of fit test

Once the *iid* classical model is rejected, the goodness of fit of the underlying data to the Yang model, needs to be examined. Hoayek et al. [31] have shown that such a goodness of fit test can be performed through examining the stochastic behavior of the inter-record time of a Yang-Nevzorov model, i.e. the time between the nth and the $(n+1)^{th}$ record defined by:

$$\Delta_{L_n} = L_{n+1} - L_n, n \geq 1.$$

Note that $L_1 = 1$, because the first observation is always considered as a record, called "trivial record." Yang [72] showed that Δ_{L_n} follows asymptotically a geometric distribution with a probability density function (which is also a distribution free property):

$$\lim_{n\to\infty} \mathbb{P}[\Delta_{L_n} = j] \equiv p_j(\gamma) = \left(1 - \frac{1}{\gamma}\right)\left(\frac{1}{\gamma}\right)^{j-1}, j \geq 1,$$

By verifying the previous distribution free property (the null hypothesis H_0), we can study the goodness-of-fit of a set of data for a Yang model. Since the previous property is asymptotic, the first observed records should be removed before the application of any goodness of fit test. Thus, in the rest of this section, N_T represents the number of records after the elimination of the first observations. The present time is denoted by T.

The considered approach consists of performing a goodness of fit test based on the well-known Chi-squared test (also written as χ^2 test) introduced by Pearson [54]. By conditioning on the event $N_T = m$ and setting K an integer ≥ 1, we consider a partition $\Pi_1, \ldots, \Pi_K$, of disjoint subsets, of $\{1, 2, \ldots, \infty\}$ with $n_k, 1 \leq k \leq K$, the number of Δ_{L_n} which fall within Π_k. $\sum_{k=1}^{K} n_k = m - 1$ is the number of inter record in a sample of m records.

By letting $\pi_k(\gamma) = \sum_{j\in\Pi_k} p_j(\gamma), 1 \leq k \leq K$, the χ^2 statistic is defined by:

$$\chi^{(\gamma)} = \sum_{k=1}^{K} (n_k - (m-1)\pi_k(\gamma))^2 / ((m-1)\pi_k(\gamma)), \tag{12}$$

where $\chi^{(\gamma)} > \chi^2_{K-1,1-\alpha}$ ($\chi^2_{K-1,1-\alpha}$ is the quantile of order $(1-\alpha)$ of the chi-square distribution with $K-1$ degrees of freedom). H_0 is rejected with an asymptotic risk level α. Then, Δ_{L_n} does not follow a geometric distribution and Yang-Nevzorov model is to be rejected.

Remark 1: While the parameter γ is unknown, the statistic $\chi^{(\gamma)}$, as defined in (12), is useless. Therefore, γ should be estimated. Based on the work of Bishop et al. [10], an estimator $\tilde{\gamma}$ of γ is obtained by minimizing $\chi^{(\gamma)}$ with respect to γ. Hence, a practically usable version of the previous statistic is:

$$\chi^{(\tilde{\gamma})} = \mathrm{argmin}\chi^{(\gamma)}. \tag{13}$$

According to Bishop et al. [10], if $\chi^{(\tilde{\gamma})} > \chi^2_{K-2,1-\alpha}$ ($\chi^2_{K-2,1-\alpha}$ is the quantile of order $(1-\alpha)$ of the chi-square distribution with $K-2$ degrees of freedom), H_0 is rejected. This puts the Yang-Nevzorov model in doubt.

Remark 2: Concerning the choice of K, this problem has been mentioned in several works [32,35,43,56]. However, none of these papers has come up with a universal solution of the problem. This has led Rayner and Rayner [58] to state that K should be fixed by considering the general context of each application.

3   Results and discussion

As a starting point, we compute the record values (R_n), the record indices (L_n) and the number of records N_T of the time series. Results are summarized in Table 2. Based on the test described in Equation (5), we check if each of the considered time series follows a classical *iid* model. The results show that both the GEOSUD and SPOT demands, fit an *iid* classical model while Landsat and Sentinel have a non-classical pattern of records (Table 3). Moreover, these results are confirmed by the goodness of fit test, where GEOSUD and SPOT represent the series having the lowest number of records; knowing that these records are concentrated among the first observations. On the other hand, Landsat and Sentinel represent the highest number of records, with an increasing rate greater than in the *iid* case.

Demand schemes							
Landsat $N_T = 9$		**Sentinel** $N_T = 7$		**SPOT** $N_T = 5$		**GEOSUD** $N_T = 5$	
Record dates	L_n	**Record dates**	L_n	**Record dates**	L_n	**Record dates**	L_n
Apr-15	1	Apr-15	1	Apr-15	1	Mar-11	1
May-15	2	May-15	2	May-15	2	May-11	3
Jun-15	3	Jun-15	3	Jun-15	3	Mar-12	13
Aug-15	5	Oct-15	7	Oct-15	7	Jun-14	40
Dec-15	9	Nov-15	8	Nov-15	8	Jan-15	47
Feb-16	11	Mar-17	24			Oct-15	56
Jan-17	22	Oct-17	31				
Mar-17	24						
Apr-17	25						

Table 2: Record results (R_n, L_n, and N_T).

Series	Landsat	Sentinel	SPOT	GEOSUD
p-value	0.25%	3.9%	23.75%	23.39%
Decision	Reject *iid*	Reject *iid*	Accept *iid*	Accept *iid*

Table 3: Goodness of fit test for the *iid* classical model with an asymptotic confidence level of 5%.

On a second step, we proceed by checking the Yang's model hypothesis for Landsat and Sentinel. We test the fitting of the inter-record time Δ_{L_n} into a geometric distribution, by using a Pearson's Chi-squared test. Table 4 shows the p-value of the $\chi^{\tilde{\gamma}}$ statistics, obtained

at an asymptotic confidence level α fixed at 5%. Recall that a p-value significantly greater than 5% indicates a good consistency of the Yang model assumptions. Otherwise, the Yang model should be rejected in order to move to a more general case, where the observations are dependent and not identically distributed. As a consequence, both time series Landsat and Sentinel have an inter-record sequence following a geometric distribution (hypothesis H_0 accepted), for which it is reasonable to adopt the Yang model.

Series	Landsat	Sentinel
p-value	0.1472	0.6868
Result	Accept H_0	Accept H_0

Table 4: Goodness of fit test for the Yang model based on Chi-squared test with an asymptotic confidence level of 5%.

In addition to the previous theoretical findings, the Landsat and Sentinel demand behavior could be reinforced through an economic reasoning:

1. First, Landsat is heading towards a greater supply of images in the coming years: Landsat_8 is adding data to its global archive at an unprecedented rate (above 700 images per day) [41, 60]. Therefore, this could bring more volatility to the market.
2. Second, future Landsat missions are designed to ensure a continuity of the oldest archive images. This feature will likely boost the usage of combined images, as Landsat is transitioning to an operational monitoring system with more explicit characterization [36, 61].
3. The huge archive accumulated of the Landsat program may probably have a strong effect on the global demand volatility, due to the increasing strategic offering of imagery [25, 26].
4. While our findings suggest that the Sentinel market presents more stability signs, a lack still remains in the interoperability and integration of this data into other information chains [30]. The computing technology and the storage power capacity are not sufficiently delivered form neither public nor private institutions in order to better cope with the new data flows [65].
5. In addition, the new generation of EO satellites from the Sentinel missions developed, is generating large amounts of data that are not easily integrated into processing chains outside the Copernicus program [59].
6. Finally, the policy of open and free access to data, which is increasingly adopted instead of the EO paying systems, could be of more influence at periods when the constraints of lack of budgets and financial models of other Earth-observation programs are questioned [7].

Moreover, as Table 5 shows, the probability of having a record can be computed for each market and for any time in the near future.

We noted that the probability of a new record on June 2018 is higher in a non-*iid* case than in the *iid* one, which fits well with our previous theoretical findings. Additionally, in line with the low chances of a having a record for all the satellite series (as shown in Table 5), no real demand record values were registered during June 2018 for the different satellite image schemes through the PEPS and GEOSUD platforms.

Based on these results, we can state the following facts:

Series	Probability of having a record on June 2018
Landsat	$\mathbb{P}[\delta_t = 1] = 14.65\%$
Sentinel	$\mathbb{P}[\delta_t = 1] = 8.60\%$
SPOT	$\mathbb{P}[\delta_t = 1] = 2.56\%$
GEOSUD	$\mathbb{P}[\delta_t = 1] = 1.14\%$

Table 5: Probability of records for each considered series.

- The Landsat and Sentinel demands follow a Yang model, with parameters $\widehat{\gamma} = 1.1713$ and 1.0907 respectively, which is evident through the relatively high probability of records. This may be seen as an indicator for an unstable demand in the future for these two schemes with a considerably high probability for future records in Landsat series.
- The SPOT series are obedient to an *iid* model, with a low probability of record due to the fact that the SPOT archive images, accessed through the PEPS operator, are no more active (the demand is based on the SPOT Word Heritage collection of images).
- The probability of a record in the GEOSUD data is close to zero. This is in line with the *iid* case and indicates that GEOSUD is stable in terms of demand. The probability of large shifts, representing a high volatility in the future, is low.

As mentioned before, the trend to full free and open data access is on the rise, even for high-resolution data. Although these data are usually distributed on a commercial basis, some are available for free via web services such as Google Earth and Bing [37, 63]. The HR satellite imagery provided through the GEOSUD SDI comes as additional concrete example. Since its launch, GEOSUD is facilitating the access to a wide variety of geospatial data and services [33]. This shift is a major result from the United States open and free data policy adopted in 2008. Other well-known systems such as CBERS[4] have followed. At the end of 2013, the European Union adopted also this strategy concerning the Sentinel Earth-observing missions. Even the "New Space" actors such as Planet Labs and SkyBox are recently making their data available for free to the academic sector and the non-public organizations [29]. The market stability in the various aspects that it can take, will be a direct consequence of these facts.

Finally, we perform an additional test to check the level of accuracy of the selected theoretical models with respect to the corresponding empirical set of observations. This step is done by comparing the expected number of records, obtained theoretically by computing $\mathbb{E}[N_T] = \sum_{t=1}^{T} P_t$, to the real values computed directly from the set of observations of each satellite image series. Results are detailed in Table 6.

Series	Landsat	Sentinel	SPOT	GEOSUD
Real number of records	9	7	5	5
Expected number of records	7.8314	6.0390	4.2279	5.0489
Exact relative error (%)	12.984	13.729	15.442	0.978

Table 6: Theoretical vs. empirical number of records.

[4]ChinaâĂŞBrazil Earth Resources Satellite program: a cooperation program between Brazil and China aiming at the development and operation of Earth observation satellites.

The real number of records is close to the expected ones, which indicates that the considered record models were well-assigned to the series of observations. In addition, the minimal error is recorded for the GEOSUD demand series, indicating a good fitting of the classical *iid* model with the GEOSUD's records behavior. Thus, an interpretation of this result relies in a stable demand expected through the GEOSUD SDI on the long term. This fact is due to the adoption of several actions, such as easy access policies to data, strong technical support for users, specific training programs, etc. As a direct consequence, GEO-SUD recorded a 7-fold increase in less than 10 years in terms of registered organizations on its platform, with more than 1000 active users.

In general, the demand shifts occurring via a SDI platform can reveal and clarify relatively the future directions to take, or the measure to deviate from in order to be well-aligned with the user's choices and orientations [11, 52]. While the platforms have become important sources of establishing EO market opportunities, the set of activities or strategies in the case of SDIs are much of importance in shaping the market dynamics. In the following, we draw on our results to suggest specific recommendations to SDIs concerning general market policy development:

Proposition 1. Diversify the SDI web portal with multi-service EO data. Given that the supply of some EO market is largely superior to other, the attractiveness of one market could be a factor to end-users to show interest in other schemes available on the SDI platform.

Proposition 2. Promote adequate and accessible SDI developments, both on upstream and downstream levels, through investment in storage capacities in order to increase the satellite image flexibility.

Proposition 3. Provide incentives and subsidies concerning pricing mechanisms in order to attract more users to the platform. The critical mass of users my increase the stability through a SDI platform.

Equipping SDIs with processing capability management (sharing of resources, cloud services, image-based applications, etc.) is necessary in enhancing the dynamic interactions occurring within the platform ecosystem. The sustainability of the future spatial data operators will rely upon the access conditions and availability of the EO resources [14, 47]. Thus, the stability of the SDI markets, which is in close relation with the development of common value-creation techniques, will depend on the implementation of open environment principles and interdependent markets and subsystems (e.g., federated user interfaces, interlinked EO data catalogues, standardized interfaces and data norms, etc.).

4   Conclusion

In general, the SDIs constitute a direct link between the large EO user community and the EO providers. This study is based on two different record models to explain the demand occurring via a SDI. The novelty results from applying the original Records theory into the geospatial science field, in order to highlight the shifts happening on various satellite image markets, each having its own characteristics.

By studying the market stability of the satellite images through a SDI platform, this paper offers an innovative and simple tool to exploit the dynamics occurring at the users' demand level. It provides an additional element in understanding the demand shifts the SDI could face on its platform. Despite the fact that the stability issue of a market is driven by several factors (price mechanisms, service quality, users' critical mass, etc.), applying the Records theory could clarify the movements occurring on a market within a time-period continuity. The market stability through SDIs depends on factors showing blurring indicators, making from the market analysis a complex exercise [42]. The challenges in the satellite imagery markets rely in the rapid evolution of technology as well as the policy frameworks of data dissemination. While the Records theory drops several constraints, it offers an additional tool to describe the market stability, and therefore, adding a precision level into the vague assumptions that could be present in a certain market. In addition, the EO market dynamics are interdependent with several technological and economic facts, that even a simple market analysis could be quasi-impossible [20]. Hence, with the unavailability of sufficient data or the complexity of a global market, the Records theory provides a simple tool allowing to group heterogeneous characteristics and implicate them together for performing a richer market analysis. As a first SDI market study using the Records theory, several ways have been revealed for improvement and extension. The main challenge relies on the further development of this research in order to assess the record values in the future. For this to be done, a more detailed theoretical work should be performed. Moreover, the methodology used in this study could be reproduced on larger scales such as the large EO observation programs (Airbus data history, NASA/Copernicus data demand history, etc.) and other types of geospatial data, where the access and demand conditions differ and the amount of data is largely greater with an impact consequently bigger. Other record models may also be included in further assessment, leading to a better examination of the demand occurring via the SDI ecosystems.

Acknowledgments

References

[1] ANDERSON, K., RYAN, B., SONNTAG, W., KAVVADA, A., AND FRIEDL, L. Earth observation in service of the 2030 Agenda for Sustainable Development. *Geo-spatial Information Science 20*, 2 (2017), 77–96. doi:10.1080/10095020.2017.1333230.

[2] ARNOLD, B. C., BALAKRISHNAN, N., AND NAGARAJA, H. N. *Records*, vol. 768. John Wiley & Sons, 2011.

[3] ARNOLD, B. C., AND VILLASEÑOR, J. A. The asymptotic distributions of sums of records. *Extremes 1*, 3 (1999), 351–363. doi:10.1023/A:1009933902505.

[4] ASCHBACHER, J., AND MILAGRO-PÉREZ, M. P. The European Earth monitoring (GMES) programme: Status and perspectives. *Remote Sensing of Environment 120* (2012), 3–8. doi:10.1016/j.rse.2011.08.028.

[5] BALINT, T. S., AND STEVENS, J. Wicked problems in space technology development at NASA. *Acta Astronautica 118* (2016), 96–108. doi:10.1016/j.actaastro.2015.09.019.

[6] BALLERINI, R., AND RESNICK, S. Records from improving populations. *Journal of Applied Probability* (1985), 487–502. doi:10.2307/3213855.

[7] BATES, J. The strategic importance of information policy for the contemporary neoliberal state: The case of Open Government Data in the United Kingdom. *Government Information Quarterly 31*, 3 (2014), 388–395. doi:10.1016/j.giq.2014.02.009.

[8] BELWARD, A. S., AND SKØIEN, J. O. Who launched what, when and why; trends in global land-cover observation capacity from civilian earth observation satellites. *ISPRS Journal of Photogrammetry and Remote Sensing 103* (2015), 115–128. doi:10.1016/j.isprsjprs.2014.03.009.

[9] BERGER, M., MORENO, J., JOHANNESSEN, J. A., LEVELT, P. F., AND HANSSEN, R. F. ESA's sentinel missions in support of Earth system science. *Remote Sensing of Environment 120* (2012), 84–90. doi:10.1016/j.rse.2011.07.023.

[10] BISHOP, Y. M., FIENBERG, S. E., AND HOLLAND, P. W. *Discrete multivariate analysis: theory and practice*. Springer Science & Business Media, New York, 2007.

[11] BORRADAILE, G. J. *Statistics of earth science data: their distribution in time, space and orientation*. Springer Science & Business Media, 2013.

[12] BOUDREAU, K., AND HAGIU, A. Platform Rules: Multi-Sided Platforms as Regulators. Tech. rep., 2008. doi:10.2139/ssrn.1269966.

[13] BUDHATHOKI, N. R., NEDOVIC-BUDIC, Z., AND OTHERS. Reconceptualizing the role of the user of spatial data infrastructure. *GeoJournal 72*, 3-4 (2008), 149–160. doi:10.1007/s10708-008-9189-x.

[14] BURROUGH, P. A., MCDONNELL, R., MCDONNELL, R. A., AND LLOYD, C. D. *Principles of geographical information systems*. Oxford university press, 2015.

[15] CHANDLER, K. N. The distribution and frequency of record values. *Journal of the Royal Statistical Society. Series B* (1952), 220–228. doi:10.1111/j.2517-6161.1952.tb00115.x.

[16] CHAPMAN, B. Waste and duplication in NASA programs: The need to enhance US space program efficiency. *Space Policy 31* (2015), 13–20. doi:10.1016/j.spacepol.2014.10.003.

[17] CLAESKENS, G., HJORT, N. L., AND OTHERS. Model selection and model averaging. *Cambridge Books* (2008).

[18] CLAVERIE, M., JU, J., MASEK, J. G., DUNGAN, J. L., VERMOTE, E. F., ROGER, J.-C., SKAKUN, S. V., AND JUSTICE, C. The Harmonized Landsat and Sentinel-2 surface reflectance data set. *Remote Sensing of Environment 219* (2018), 145–161. doi:10.1016/j.rse.2018.09.002.

[19] DE MONTALVO, U. W. *Mapping the determinants of spatial data sharing*. Routledge, 2017.

[20] DENIS, G., DE BOISSEZON, H., HOSFORD, S., PASCO, X., MONTFORT, B., AND RANERA, F. The evolution of Earth Observation satellites in Europe and its impact on the performance of emergency response services. *Acta Astronautica 127* (2016), 619–633. doi:10.1016/j.actaastro.2016.06.012.

[21] DONLON, C., BERRUTI, B., BUONGIORNO, A., FERREIRA, M.-H., FÉMÉNIAS, P., FRERICK, J., GORYL, P., KLEIN, U., LAUR, H., MAVROCORDATOS, C., AND OTHERS. The global monitoring for environment and security (GMES) sentinel-3 mission. *Remote Sensing of Environment 120* (2012), 37–57. doi:10.1016/j.rse.2011.07.024.

[22] FORESMAN, T. W. Evolution and implementation of the Digital Earth vision, technology and society. *International journal of digital earth 1*, 1 (2008), 4–16. doi:10.1080/17538940701782502.

[23] GAWER, A., AND CUSUMANO, M. A. How companies become platform leaders. *MIT Sloan management review 49*, 2 (2008), 28.

[24] GEOSPATIAL INDUSTRY OUTLOOK AND READINESS INDEX, G. Geobuiz. Tech. rep., Accessed on 10 february, 2019, 2018.

[25] GOWARD, S., ARVIDSON, T., WILLIAMS, D., FAUNDEEN, J., IRONS, J., AND FRANKS, S. Historical record of Landsat global coverage. *Photogrammetric Engineering & Remote Sensing 72*, 10 (2006), 1155–1169. doi:10.14358/PERS.72.10.1155.

[26] GUO, H., LIU, Z., JIANG, H., WANG, C., LIU, J., AND LIANG, D. Big Earth Data: A new challenge and opportunity for Digital Earth's development. *International Journal of Digital Earth 10*, 1 (2017), 1–12. doi:10.1080/17538947.2016.1264490.

[27] HANSEN, M. C., AND LOVELAND, T. R. A review of large area monitoring of land cover change using Landsat data. *Remote sensing of Environment 122* (2012), 66–74. doi:10.1016/j.rse.2011.08.024.

[28] HARRELL, F. E. Regression modeling strategies. *as implemented in R package âĂŸrms' version 3*, 3 (2014).

[29] HARRIS, R., AND BAUMANN, I. Open data policies and satellite Earth observation. *Space Policy 32* (2015), 44–53. doi:10.1016/j.spacepol.2015.01.001.

[30] HELDER, D., MARKHAM, B., MORFITT, R., STOREY, J., BARSI, J., GASCON, F., CLERC, S., LAFRANCE, B., MASEK, J., ROY, D., AND OTHERS. Observations and Recommendations for the Calibration of Landsat 8 OLI and Sentinel 2 MSI for improved data interoperability. *Remote Sensing 10*, 9 (2018), 1340. doi:10.3390/rs10091340.

[31] HOAYEK, A. S., DUCHARME, G. R., AND KHRAIBANI, Z. Distribution-free inference in record series. *Extremes 20*, 3 (2017), 585–603. doi:10.1007/s10687-017-0283-7.

[32] INGLOT, T., LEDWINA, T., AND OTHERS. Asymptotic optimality of data-driven Neyman's tests for uniformity. *The Annals of Statistics 24*, 5 (1996), 1982–2019. doi:10.1214/aos/1069362306.

[33] JABBOUR, C., HOAYEK, A., MAUREL, P., REY-VALETTE, H., AND SALLES, J.-M. How Much Would You Pay for a Satellite Image?: Lessons Learned From French Spatial-Data Infrastructure. *IEEE geoscience and remote sensing magazine 8*, 4 (2020), 8–22. doi:10.1109/MGRS.2019.2941751.

[34] JABBOUR, C., REY-VALETTE, H., MAUREL, P., AND SALLES, J.-M. Spatial data infrastructure management: A two-sided market approach for strategic reflections. *International Journal of Information Management 45*, 69-82 (2019). doi:10.1016/j.ijinfomgt.2018.10.022.

[35] KALLENBERG, W. C. M., AND OTHERS. On moderate and large deviations in multinomial distributions. *The Annals of Statistics 13*, 4 (1985), 1554–1580. doi:10.1214/aos/1176349755.

[36] KARDAN, N., DABNEY, P., AND BABU, S. Landsat Missions to Sustainable Land Imaging Technology Program. In *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium* (2018), IEEE, pp. 6336–6337. doi:10.1109/IGARSS.2018.8517943.

[37] KEYSERS, J. *Review of digital globes 2015.* Australia and New Zealand Cooperative Research Centre for Spatial Information, 2015.

[38] KHRAIBANI, Z., JACOB, C., DUCROT, C., CHARRAS-GARRIDO, M., AND SALA, C. A Non Parametric Exact Test Based on the Number of Records for an Early Detection of Emerging Events: Illustration in Epidemiology. *Communications in Statistics-Theory and Methods 44* (2015), 726–749. doi:10.1080/03610926.2013.799695.

[39] KRUSE, J. B., CROMPVOETS, J., AND PEARLMAN, F. *GEOValue: The Socioeconomic Value of Geospatial Information.* CRC Press, 2017.

[40] LEE, C.-F., AND LEE, J. C. *Handbook of financial econometrics and statistics.* Springer, 2015.

[41] LOVELAND, T. R., AND IRONS, J. R. Landsat 8: The plans, the reality, and the legacy. *Remote Sensing of Environment 185* (2016), 1–6. doi:10.1016/j.rse.2016.07.033.

[42] MACHARIS, C., AND CROMPVOETS, J. A stakeholder-based assessment framework applied to evaluate development scenarios for the spatial data infrastructure for Flanders. *Computers, Environment and Urban Systems 46* (2014), 45–56. doi:10.1016/j.compenvurbsys.2014.04.001.

[43] MANN, H. B., AND WALD, A. On the choice of the number of class intervals in the application of the chi square test. *The Annals of Mathematical Statistics 13*, 3 (1942), 306–317.

[44] MCDOUGALL, K. The potential of citizen volunteered spatial information for building SDI.

[45] MILLER, H. M., RICHARDSON, L., KOONTZ, S. R., LOOMIS, J., AND KOONTZ, L. Users, uses, and value of Landsat satellite imageryâĂŤResults from the 2012 survey of users. *US Geological Survey Open-File Report 1269*, 2013 (2013), 51.

[46] MULLAINATHAN, S., AND SPIESS, J. Machine learning: an applied econometric approach. *Journal of Economic Perspectives 31*, 2 (2017), 87–106. doi:10.1257/jep.31.2.87.

[47] NATIVI, S., MAZZETTI, P., SANTORO, M., PAPESCHI, F., CRAGLIA, M., AND OCHIAI, O. Big data challenges in building the global earth observation system of systems. *Environmental Modelling & Software 68* (2015), 1–26. doi:10.1016/j.envsoft.2015.01.017.

[48] NEVZOROV, V. B. Records for nonidentically distributed random variables. *Proceedings of the Fifth Vilnius Conference 2* (1990), 227–233.

[49] NEVZOROV, V. B. *Records: mathematical theory*. American Mathematical Society, Rhode Island, 2001.

[50] NOSAVAN, J., MOREAU, A., AND HENRY, P. Spot world heritage: exploring the past. In *Sensors, Systems, and Next-Generation Satellites XXII* (2018), vol. 10785, International Society for Optics and Photonics, p. 107850T. doi:10.1117/12.2324675.

[51] PARKER, G., AND VAN ALSTYNE, M. Innovation, openness, and platform control. *Management Science* (2017). doi:10.1287/mnsc.2017.2757.

[52] PEARLMAN, F., PEARLMAN, J., BERNKNOPF, R., COOTE, A., CRAGLIA, M., FRIEDL, L., GALLO, J., HERTZFELD, H., JOLLY, C., MACAULEY, M., AND OTHERS. Assessing the socioeconomic impact and value of open geospatial information. Tech. rep., US Geological Survey, 2016.

[53] PELTON, J. N., MADRY, S., AND CAMACHO-LARA, S. *Handbook of satellite applications*. Springer Publishing Company, Incorporated, 2017.

[54] PLACKETT, R. L. Karl Pearson and the chi-squared test. *International Statistical Review/Revue Internationale de Statistique* (1983), 59–72. doi:10.2307/1402731.

[55] PWC. Study to examine the socio- economic impact of Copernicus in the EU. Tech. rep., 2016.

[56] QUESENBERRY, C. P., AND MILLER JR, F. L. Power studies of some tests for uniformity. *Journal of Statistical Computation and Simulation 5*, 3 (1977), 169–191. doi:10.1080/00949657708810150.

[57] RAJABIFARD, A., AND OTHERS. Data integration and interoperability of systems and data. In *2nd Preparatory Meeting of the Proposed UN Committee on Global Geographic Information Management* (2010).

[58] RAYNER, G. D., AND RAYNER, J. C. W. Power of the Neyman smooth tests for the uniform distribution. *Advances in Decision Sciences 5*, 3 (2001), 181–191. doi:10.1207/S15327612JAMD0503_3.

[59] REGAN, A., SILVESTRIN, P., AND FERNANDEZ, D. Sentinel Convoy: Synergetic Earth Observation with Satellites Flying in Formation with European Operational Missions. In *Living Planet Symposium* (2016), vol. 740, p. 379.

[60] ROY, D. P., WULDER, M. A., LOVELAND, T. R., WOODCOCK, C. E., ALLEN, R. G., ANDERSON, M. C., HELDER, D., IRONS, J. R., JOHNSON, D. M., KENNEDY, R., AND OTHERS. Landsat-8: Science and product vision for terrestrial global change research. *Remote sensing of Environment 145* (2014), 154–172. doi:10.1016/j.rse.2014.02.001.

[61] SCHROEDER, T. A., LOVELAND, T., WULDER, M. A., AND IRONS, J. R. Landsat Science Team: 2017 Winter Meeting Summary. *The Earth Observer 28*, 6 (2017), 45–50.

[62] SHELESTOV, A., LAVRENIUK, M., KUSSUL, N., NOVIKOV, A., AND SKAKUN, S. Exploring Google earth engine platform for Big Data Processing: Classification of multi-temporal satellite imagery for crop mapping. *Frontiers in Earth Science 5* (2017), 17. doi:10.3389/feart.2017.00017.

[63] SWAIN, N. R., LATU, K., CHRISTENSEN, S. D., JONES, N. L., NELSON, E. J., AMES, D. P., AND WILLIAMS, G. P. A review of open source software solutions for developing water resources web applications. *Environmental Modelling & Software 67* (2015), 108–117. doi:10.1016/j.envsoft.2015.01.014.

[64] TORRES, R., SNOEIJ, P., GEUDTNER, D., BIBBY, D., DAVIDSON, M., ATTEMA, E., POTIN, P., ROMMEN, B., FLOURY, N., BROWN, M., AND OTHERS. GMES Sentinel-1 mission. *Remote Sensing of Environment 120* (2012), 9–24. doi:10.1016/j.rse.2011.05.028.

[65] WICKS, D., JONES, T., AND ROSSI, C. Testing the Interoperability of Sentinel 1 Analysis Ready Data Over the United Kingdom. In *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium* (2018), IEEE, pp. 8655–8658. doi:10.1109/IGARSS.2018.8518120.

[66] WILLMES, C., BECKER, D., VERHEUL, J., YENER, Y., ZICKEL, M., BOLTEN, A., BUBENZER, O., AND BARETH, G. PaleoMaps: SDI for open paleoenvironmental GIS data. *International Journal 12* (2017), 39–61. doi:10.2902/1725-0463.2017.12.art3.

[67] WITTEN, I. H., FRANK, E., HALL, M. A., AND PAL, C. J. *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, 2016.

[68] WOODCOCK, C. E., ALLEN, R., ANDERSON, M., BELWARD, A., BINDSCHADLER, R., COHEN, W., GAO, F., GOWARD, S. N., HELDER, D., HELMER, E., AND OTHERS. Free access to Landsat imagery. *Science 320*, 5879 (2008), 1011.

[69] WULDER, M. A., MASEK, J. G., COHEN, W. B., LOVELAND, T. R., AND WOODCOCK, C. E. Opening the archive: How free data has enabled the science and monitoring promise of Landsat. *Remote Sensing of Environment 122* (2012), 2–10. doi:10.1016/j.rse.2012.01.010.

[70] WULDER, M. A., WHITE, J. C., LOVELAND, T. R., WOODCOCK, C. E., BELWARD, A. S., COHEN, W. B., FOSNIGHT, E. A., SHAW, J., MASEK, J. G., AND ROY, D. P. The global Landsat archive: Status, consolidation, and direction. *Remote Sensing of Environment 185* (2016), 271–283. doi:10.1016/j.rse.2015.11.032.

[71] YANG, C. P., CAO, Y., EVANS, J., KAFATOS, M., AND BAMBACUS, M. Spatial Web portal for building spatial data infrastructure. *Geographic Information Sciences 12*, 1 (2006), 38–43. doi:10.1080/10824000609480616.

[72] YANG, M. C. K. On the distribution of the inter-record times in an increasing popula-
     tion. *Journal of Applied Probability* (1975), 148–154. doi:10.2307/3212417.