Article Type: Research Paper

Big-Data Analytics Framework For Incorporating Smallholders In Sustainable Palm Oil Production

Manish Shukla\textsuperscript{a}, and Manoj Kumar Tiwari\textsuperscript{b}

\textsuperscript{a} Durham University Business School, Durham University, Durham, UK, DH1 3LB
\textsuperscript{b} Department of Industrial and Systems Engineering, Indian Institute of Technology, Kharagpur 721302, India

Abstract:
This paper aims to address the constraints faced in incorporating smallholders in sustainable palm oil production. There exists literature that acknowledges the need for incorporating smallholders in the production of sustainable palm oil but none has proposed a solution beyond ‘Roundtable on Sustainable Palm Oil’ (RSPO) certification. In the current business scenario, several organizations are struggling to procure RSPO certified palm oil even after committing huge resources. RSPO, though a good first step, has a major process and capacity constraints resulting in long processing times, delays, and lack of traceability for the customers. This paper proposes a Big Data Analytics framework enabled by cutting-edge technologies to incorporate smallholders in the RSPO certification process. The data used was collected through farm visits, stakeholder meetings, key stakeholder interviews, and, secondary sources. The proposed framework not only addresses the limitation of the current certification process but also converts it from being punitive to preventive. The outcomes of this research will be extremely useful for all the stakeholders in the palm oil supply chain.

Keywords: Big Data Analytics, Sustainable Production, Palm Oil, Roundtable on Sustainable Palm Oil, New Technologies

\textsuperscript{1} Corresponding author. email: manish.shukla@durham.ac.uk
1. Introduction:

Over the years, several large organizations such as Unilever, Nestle, Pepsi Co, Procter & Gamble, Ferrero, etc. have been accused by the Non-Governmental Organizations (NGOs) for using unsustainable palm oil in their products. NGOs are pressurising the companies to procure sustainable palm oil that is not produced on a deforested land. Of all the possible proposed solutions, procuring ‘Roundtable on Sustainable Palm Oil’ (RSPO) certified oil is the most widely accepted solution by the industry. Several organizations have made a public pledge to procure 100% RSPO certified palm oil by a certain date to eliminate deforestation from the palm oil production. Other companies are also queuing-up to take the similar path. But, prominent NGOs such as Greenpeace and others are dissatisfied with this option, as they believe that the current RSPO certification is ineffective to eliminate deforestation. In a report titled “Certifying Destruction” Greenpeace has provided compelling reasons to look beyond RSPO.

Most of the large multinational companies procure the palm oil from short-term contracts/joint ventures with palm oil mills in Indonesia and Malaysia. The palm oil mills procure the palm fruit branches from their own smallholder, large estates, and traders. The traders procure the palm fruit branches from small estates and independent smallholders. The palm oil mills mostly control the upstream operations whereas the large multinational companies have least information and control of the upstream supply chain. Thus, in the current scenario, it is almost impossible for these companies to move beyond the industry accepted RSPO certification even if they are pressurised by the NGOs. At this stage, there is no doubt that RSPO is a “Good First Step” to move towards sustainability in the palm oil production. But a closer look at the RSPO certification process reveals that the entire process has significant resource and administrative limitations. There is certainly a need to look not only beyond but also within the current certification system to make it more efficient and transparent.

2 http://www.greenpeace.org/international/en/publications/Campaign-reports/Forests-Reports/Certifying-Destruction/
RSPO has 8 principles and each principle has several criteria for which the documentations, inspection, or evidence of appropriate training is required for validation. Overall the criteria when added together require more than 140 data points for validation. The planters (estates and smallholder) are expected to ensure the availability of these data points at the time of audit. Organized smallholders and estates find it less problematic due to the availability of support and certainty. But it is highly difficult for independent smallholders to manage the additional administrative tasks while doing their regular farming activates, as the overall process is very time consuming and costly. Moreover, the method and audit of the RSPO certification is highly questionable as there is no way to evaluate the certification process or quality of the audit. In addition, it is to be noted that entire auditing exercise is ex-post, which essentially means that if a smallholder is engaged in unsustainable practices, the auditors may not be able to prevent it but will not certify the plantation in the next audit. Thus, the need for huge data, administrative burden, lack of transparency, and nature of the RSPO certification motivates us to look for alternative solutions that can enhance independent smallholder participation in the certification process.

A critical examination of the RSPO criteria for certification reveals that data for most of the criteria can be gathered in real time from alternate source without any extra efforts from the smallholders. Advance technologies such as global positioning system (GPS), satellite image processing, sensors, radio frequency identification (RFID) tags, etc. can be used to gather real-time data and make the process efficient and transparent. Thus, we propose a big data analytics framework enabled by advance technologies to incorporate smallholders in the RSPO certification process. The data used to develop the framework was collected through farm visits, stakeholder meetings, key-stakeholder interviews and secondary sources. It is expected that the proposed big data analytics framework will reduce smallholders’ burden for data management, reduce the dependence on the third party auditors, reduce the cost of certification, and make the buyers more confident about the product. It is also expected that the proposed framework will not only addresses the limitation of the current certification process but will also make it preventive in nature. The outcomes of
this research will be extremely useful for all the stakeholders in the palm oil supply chain. The novelty of this research is that:

- It conducts an in-depth analysis of the current certification process
- Highlights the practical limitations of RSPO
- Proposes a solution to overcome the identified limitations.

Rest of the paper is structured as follows: Section 2 presents a detailed literature review addressing sustainable supply chains and use of technology. The research methodology is presented in section 3. Section 4 presents the problem context of palm oil production. The analysis of the current certification process is presented in section 5 and the big-data analytics framework is presented in section 6. Section 7 presents results and discussion. The paper is concluded in section 8 with a discussion about scope for future research.

2. Literature Review:

Sustainable supply chain management has emerged as the prime concern for the practitioners and academicians in the recent years (Seuring and Müller, 2008; Barber, Beach, and Zolkiewski, 2012). Even though there are differences among the researchers about the definition of sustainability (Carter and Rogers, 2008), it may be understood as “development that meets the needs of the present without compromising the ability of future generations to meet their own needs” (Brundtland, 1985). Sustainability is more than just the environmental or social issue rather is it at the intersection of the economical, environmental, and social issues (Carter and Easton, 2011), also know as Triple Bottom Line (Elkington, 1997). Over the years, sustainability has emerged for being an operational concern to being the most important strategic issue for the organizations. According to Carter and Easton (2011), sustainability has become a mandatory requirement for the organizations in the 21st century.

Researchers found that sustainable supply chain initiatives have a positive impact on environmental, economic, and intangible outcomes (Eltayeb, Zailani, and Ramayah, 2011). Several efforts are being made to identify the essential elements that need to be taken into account in order to incorporate sustainability in a supply chain (Pagell and Wu, 2009; Chaabane, Ramudhin, and Paquet, 2011; Marshall et al., 2015). One of the prime elements is managing
buyer-supplier relationships because sustainable initiatives generally tend to move toward long-term partnerships (Seuring and Müller, 2008; Dou, Zhu, and Sarkis, 2014). With the change in relationship, there is also a need to re-visit the incentive alignment and performance indicators (Seuring and Müller, 2008). Moreover, sustainability creates a need to look beyond the immediate relationships and collaborate across the supply chain as each player has to share the responsibility and coordination can create a competitive advantage (Seuring and Müller, 2008; Gold, Seuring, and Beske, 2010; Macchion et al., 2016). This is a major challenge for the multi-national organization having global supply chain operations. This challenge becomes more difficult in cases where there are a large number of small suppliers at the upstream of the chain, for example coffee, banana, or palm oil supply chains.

Agricultural produce such as rubber, cotton, cocoa, palm oil, etc. is an essential raw material for several industrial products. Several large companies and retailers are involved in selling fresh or processed food products. Over the years, consumers have become increasingly concerned about the origin and production conditions of their food and non-food items and often emphasis on reliable evidence for traceability (Opara, 2003). Studies have found that sustainable practices are highly influenced by the customer/consumers (Glover et al., 2014).

In addition to the consumer pressure, the stringent legal regulations require organizations to follow sustainable practices across their supply chain (Chaabane, Ramudhin, and Paquet, 2011; Hitchcock, 2012). There are also examples where NGOs such as Greenpeace have created significant pressure on several organizations such as Nestle and others for using unsustainable ingredients (palm oil) in their products (Wolf, 2014). These challenges are often taken care by incorporating partnered governance international certification schemes such as fair-trade, Roundtable on Sustainable Palm Oil (RSPO), etc. These certification schemes are supposed to addresses the consumer concerns for tractability and to provide operational standards for the producers (Nagiah and Azmi, 2013). Thus, these efforts to enable sustainable practices across the supply chains have to incorporate the upstream suppliers such as the smallholders and small agri-food organizations.
Researchers have studied the challenges of incorporating upstream suppliers and proposed solutions to develop supplier involvement, performance management, and collaboration (Gold, Seuring, and Beske, 2010; Dou, Zhu, and Sarkis, 2014). Researchers found that the approach to achieve sustainability varies based on the interest of the key supply chain players (MacCarthy and Jayarathne, 2012). They also found an increase in interest of the stakeholders towards social and environmental factors in decision-making while economic factors still have the highest importance (Vasileiou and Morris, 2006). Still the smallholders and small agri-food organizations are finding it extremely difficult to collaborate with the large organizations due to issues such as the ambiguity of certification process, the reliability of different certifications, the changing dynamics of buyer-supplier relationship, and, the lack of means to ensure sustainability to the final consumers. Researchers have explored the enablers and constraints faced by the small companies while implementing sustainable practices. Some of the prevalent enablers include the improvements in livelihood, higher premiums, better yields, more market opportunities for the smallholders. For example, Kilian et al. (2006) compared the differences in costs and premium prices for the organic and fair-trade coffee category to identify the benefits of the certification schemes for the smallholders. Fayet and Vermeulen (2014) based on their investigation of nine cases of cotton supply chain in India reported that there are improvements in livelihood of the smallholders. On the other hand, Bacon et al. (2008) reported that though the sustainable coffee certification schemes for smallholders in Nicaragua has an advantage in few areas but in general has minimal effect on income. Therefore, even though there are benefits of the sustainable practices for the small holder, these are uncertain and ambiguous.

The complexity of implementing sustainable practices and uncertainty of the incentives often limit the smallholders participation. Several researchers have explored this issue for example, MacDonald (2007) discussed the complexities of incorporating sustainability in coffee supply chains and highlighted the difficulties faced by the small and marginal farmers. Bitzer, Francken, and Glasbergen (2008) addressed the difficulties and limitations of inter-sectoral partnership to achieve sustainability taking a case of coffee supply chains. De
Brito, Carbone, and Blanquart (2008) studied the fashion retail supply chain and explored the challenges faced by the upstream-stakeholders producing the raw material.

Researchers have also proposed potential solutions to address the challenges in incorporating sustainable practices in agri-food supply chains. For example, Svensson and Wagner (2012) proposed a nine-stage implementation plan for a commodity (dairy) product. Yakovleva (2007) used a sustainability assessment model to investigate the chicken and potato supply chains in UK. Sureeyatanapas, Yang, and Bamford (2015) explored the challenges of incorporating sustainability in the sugar industry. Bourlakis et al. (2014) examined the Greek dairy industry for sustainable practices and reported that there is a need for improvement in key sustainability performance indicators. In case of palm oil, organizations tend towards adopting RSPO guidelines for achieving sustainable production practices. RSPO is an internationally recognised certification scheme to enable sustainable palm oil production. It is a key example of partnered governance where all the stakeholders participate in defining the guidelines and the governance structure. Schouten and Glasbergen, (2011) investigated the legitimacy of RSPO by analysing its legal, moral, and, acceptance aspects as it interacts with a large number of stakeholder and local and international legislations. RSPO though an acceptable certification, still has several major limitations and scope for improvements (Nikoloyuk, Burns, and de Man, 2010). One of the major limitations may be the cost and documentation requirement. Lack of resources and unavailability of information inhibits smallholders’ participation in the certification schemes. It becomes highly challenging for the large organizations to incorporate smallholders in their sustainable supply chains (Nagiah and Azmi, 2013).

The cost of certification can be managed up to a certain level by aligning the incentive systems, participation by the large organizations, financial institutions, and, local governments. But, the bottlenecks in the certification process and the challenge of ensuring the compliance to the certification process through error free data and documents are still a major roadblock. Some of these challenges may be taken care by re-visiting the certification process from a business process.
perspective to identify and eliminate the existing bottlenecks. The challenges of availability and reliability may be addressed by incorporating the emerging advance technologies for data collection and analysis. Advance technology such as RFID, sensors, web-based solutions, Internet of Things (IOT) can be used to redesign the process, improve data collection, and, enhance visibility across the supply chain (Pero and Rossi, 2014; Ferretti and Schiavone, 2016; Shamsuzzoha et al., 2016). Several researchers have used data from conventional and non-conventional sources such as (tweets, videos, post, etc.) to enable sustainable practices. For example, Papadopoulos et al. (2017) proposed and validated a theoretical framework using big data for sustainability of supply chain networks. Zhao et al. (2017) proposed a multi-objective optimization model for green supply chain management using big data analysis. The existing technologies can generate a vast amount of data that can facilitate value creation and create a competitive advantage for the organizations (Tan et al., 2015). There exists literature that discusses the use of existing technology such as phones, cameras, sensors, satellites (Seelan et al., 2003; Mulla, 2013) and emerging technology such as unmanned aerial vehicles or drones (Herwitz, 2004; Everaerts, 2008; Zhang and Kovacs, 2012) for agriculture production. These can be used to reduce data collection burden and also to ensure higher reliability of the collected data. Moreover, these can be used to conduct an in-depth analysis and enable the auditors in the certification schemes to analyse and interpret the data. Though there is immense potential benefits for analysing the current certification process and using advance technology to overcome the bottlenecks, there exists no literature that attempts to address this concerns. Thus, this research aims to address this gap in the literature.

3. Research Methodology
This section presents the methodology adopted for this research. As this research is exploratory in nature, thus a case study based approach was used to understand the context, identify the problem, and find the potential solutions. We used the approach for theory building from case studies proposed by Eisenhardt (1989). Figure 1 presents the details of the research methodology applied for this paper.
The first step in the approach is 'Getting started' where the activity is to define the research question or a priori construct. In this case, the research question was to investigate “how to incorporate smallholders in sustainable production". The second step required us to select the cases. Case study selection is a very important and critical decision. Eisenhardt (1989) discussed the importance of case selection and its impact on generalizability of the results. The view is to select cases such that the emergent theory can be either tested or extended. Seawright and Gerring (2008) presented a detailed discussion about case selection methods. They discussed seven types of cases studies from typical to most different. Gibbert, Ruigrok, and Wicki (2008) argued that in case selection internal and construct validity should not be ignored to ensure external validity. They emphasized on the logic, reasoning, and, transparency of the overall process. We identified RSPO certification process for smallholders as a case to be further analysed. This was because it represents a typical situation for smallholders, which is common in other sustainable certification schemes such as fair trade. There are several other similar situations such as coffee certification, banana certification, etc. But, the difficulties of incorporating the smallholders in sustainable production are same for all cases irrespective of the produce. Thus the selected case is a good representation of the entire population. The internal and construct validity was maintained by following the process suggested by Eisenhardt (1989). Step 3 of the approach requires crafting the instruments and protocols. It encourages identification of multiple data sources, developing interview questions, planning field visits and collection of data from the secondary sources. At this step, data from secondary sources such as academic literature, reports from Greenpeace, and, websites of RSPO, etc. was analysed to develop a basic understanding of existing problem of sustainable palm oil production. This analysis also provided basic information about the stakeholders (buyers, mills, smallholders, NGOs, certification bodies, auditors, etc.), RSPO certification process, and the role of smallholders in the overall scenario. Based on this understanding, we conducted a detailed literature review to explore the drivers, measures, and, challenges in implementing sustainable production practices across various agriculture produce with a focus towards
smallholders. Based on the literature review, we were able to identify the existing gap in the literature.

The next step of the approach is about entering the filed and collecting data. We conducted initial field visits and had open-ended interviews with the smallholders, consolidators, dealers, and NGOs (working for smallholders). This enabled us to develop a better understanding of the certification issue from the smallholders’ perspective. Data about the RSPO certification process was collected from the RSPO website, stakeholders (smallholders, consolidators, and estate managers) and experts (NGO involved in implementation of RSPO certification).

In the next step we started analysing the data to identify trends and patterns. We mapped the current RSPO certification process based on the data and key expert
interview. Initially the certification process was mapped based on the secondary sources and was shared with the experts for verification. Based on their suggestions, the process mapping was further developed and validated. This provided us with an in-depth and realistic understanding of the process that is not otherwise visible to the downstream players. It is also to be noted that the real certification process map for smallholder RSPO certification is also not available in the existing literature. Once the process was mapped, we started analysing the process in terms of time, data, and resources required. We also analysed the responsibility of different actors involved in the certification process. Based on the process analysis, we were able to identify the key bottlenecks in the process in terms of the resources requirements, delays, lack of transparency, and, audit quality. We identified the major data points that are required in the audit process, their sources, and, responsible actors. This enabled the formulation of the hypothesis that ‘a Big-data analytics framework can overcome the limitations of the current certification process’. We analysed the literature to identify the alternate means to source the required data using advance technology. Based on the outcomes, we proposed a big data analytics framework and discussed its characteristics, implementations, and, limitations.

4. Sustainable palm oil production by smallholders: A case study

Smallholders are defined as farmers with palm plantations anywhere from 5 Hectare to 40 Hectare (Ismail, Simeh, and Noor, 2003; Cramb and Sujang, 2013). Among the smallholders, there are organized farmers, promoted by government scheme and linked to a specific mill, especially in Malaysia (Ismail, Simeh, and Noor, 2003; Cramb and Sujang, 2013). The other kind of smallholders is independent farmers, producing and selling palm oil through consolidators and traders. These independent farmers are at high risk for sustainable palm oil production given their relatively smaller size and absence of a binding obligation. Additionally, their large numbers, geographical spread, and, absence of leadership drastically surges the transaction cost to integrate them. This inhibits large organizations to consider them as a viable source of sustainable palm oil taking into account the associated risk and transaction cost. The organizations generally want to curtail the risk associated with any potential
malpractice, so they find it convenient to procure from large estates and organized smallholders who can be held accountable. This results in smallholders, especially independent, being ignored by the large organizations’ in their efforts to respond to the calls for sustainable palm oil production. For the large organizations’ this may be a secure way to ascertain their claims and to satisfy stakeholders’ quest for sustainability. But this isolates the smallholders, escalating the risk for unsustainable practices. Because if they are excluded from the organized procurement process citing absence of RSPO certificate, they may be forced to trade in open markets where buyers may be eager to compromise sustainability over cost. For examples, the household buyers in countries such as India and China, where palm oil is consumed as cooking oil, may not be aware or educated about the unsustainable practices and may be willing to buy the oil if offered at a cheaper price. Therefore, though the large organizations may be able to wash-off their hands and come clean, the problem of unsustainable practices may still exists only with different set of producers and buyers who might be more challenging to deal with.

The numbers that we are arguing is not marginal, around 3 million smallholder are involved in palm oil plantations, contributing around 40% of the total world production (Vermeulen and Goad, 2006). Either it be the major palm oil producing countries such as Indonesia and Malaysia or minor players such as Thailand and Nigeria, smallholders contribution is substantial enough to impact the national image and global environment (Vermeulen and Goad, 2006). It will be a grave mistake to exclude the smallholders from the sustainable palm oil supply chains and push them towards open markets. They shall be encouraged to participate in RSPO certification to not only ripe the benefits of the sustainable certificates, but also to get the hidden advantages of improved yields, slashed water and fertiliser quantity, and, reach to buyers. RSPO has only recently started encouraging smallholders and have introduced a certification scheme for them.

4.1. RSPO Certification For Smallholders

The RSPO certification process is a bit different for the smallholders as compared to the other farmers and estates. There is a need for a lot of groundwork before a smallholder is ready for certification. Due to the small size and high administrative
costs, the smallholders are grouped together for certification process. In general, the group needs external support and training for the initial period to reach a satisfactory level for certification. Most of the times, this support is extended by the NGOs working in the rural areas. These NGOs teach the local-youth to become the certification advisor for the group. The NGOs with the help of local-youth advisors and volunteers start educating the smallholder groups about various requirements and farming practices to be adopted. Once the smallholder group is ready for the certification they need to contact the RSPO audit team for certification and have to put a one-month public notice declaring their intention for RSPO certification.

Figure 2: RSPO Certification Process for Smallholders
After the end of notice period, the RSPO team visits the plantation to conduct survey and inspection. This is in general a 2 day visit, where on the first day, the audit team interviews the plantation manager (NGOs and local-youth) about the documentation, records, etc. and on the second day they interview the smallholders and conduct physical inspection of the plantations. Based on the evidence collected from the two-day visit, the audit team proposes their recommendation to the RSPO. If the recommendations are favourable then the certification body awards a five-year certificate for sustainable production to the smallholders, which needs to be annually reviewed. Detailed steps of the smallholder certification process are presented in Figure 2.

5. Current RSPO Certification Process

The RSPO certification process has three key elements namely: standards, accreditation, and process requirements. The standards define the baseline requirements for certifying any stakeholder. These standards are composed of a set of principles and criteria that ensure sustainable palm oil production. These principles and criteria are a guiding document that has different national interpretations due to the difference in legal, social, and, cultural eco-system across nations. This enables fine integration of national legislations and stakeholders with global standards. These principles and criteria are critically analysed and evaluated in every five years. There are eight guiding principles for RSPO namely, transparency, compliance to legislation, economic viability, best practices, environment responsibility, social accountability, responsible development of new plantation, and, continuous improvement. Each of these principles is quantified by a set of criteria. In total, these 8 principles are assessed by more than 40 criteria, which have more than 140 indicators. These indicators require minimum one evidence for certification. For example, in case of transparency principle, criteria include online availability of documents, record for ethical practice, etc. indicator will include records, documentation, written policy for ethical conduct, which require several documents to support. To simply for the readers, the RSPO certification process requires evidence for 140 factors, with 1 or more supporting evidence for each factor. From a data collection perspective, the RSPO certification process requires data from around 140

[^3]: http://www.rspo.org/certification/how-rspo-certification-works
data sources that includes, land records, policy documents, training certificates, plantation maps, etc. In case of RSPO certification, these documents have to be collected and complied by the smallholders to meet the certification requirements. Table 1 presents the details of the principles, criteria, indicators, and, evidence requirements for the RSPO certification process for smallholders.

<table>
<thead>
<tr>
<th>Principles</th>
<th>Criteria</th>
<th>Requirements</th>
<th>Ownership</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Commitment to transparency</td>
<td>Land Records, Policy documents, etc.</td>
<td>Legal and Policy Documents</td>
<td>Smallholder</td>
</tr>
<tr>
<td>2. Compliance with applicable laws and regulations</td>
<td>MPOB License, Driving permit, etc.</td>
<td>License and permits</td>
<td>Smallholder</td>
</tr>
<tr>
<td>3. Commitment to long term economic and financial viability</td>
<td>Not applicable to smallholders</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>4. Uses of appropriate best practices by growers and millers</td>
<td>Good agricultural practices, soil health and water level management</td>
<td>Training</td>
<td>Smallholder</td>
</tr>
<tr>
<td>5. Environmental responsibility and conservation of natural resources and biodiversity</td>
<td>HCV conservation, waste handling, land preparation, etc.</td>
<td>Inspection</td>
<td>Smallholder</td>
</tr>
<tr>
<td>6. Responsible consideration of employees, and of individuals and communities affected by growers and millers</td>
<td>Labour laws, Child right acts, etc.</td>
<td>Inspection</td>
<td>Smallholder</td>
</tr>
<tr>
<td>7. Responsible development of new plantings</td>
<td>Legal rights, HCV conservation, Burning, GHG emission</td>
<td>Inspection and training</td>
<td>Smallholder</td>
</tr>
<tr>
<td>8. Commitment to continuous improvement in key areas of activity</td>
<td>Yield improvement, agricultural practice improvements</td>
<td>Training</td>
<td>Smallholder</td>
</tr>
</tbody>
</table>

Table 1: Details of the principles, criteria, documents and ownership for RSPO certification

It can be appreciated that this is a comprehensive list of evidence, which may be extremely difficult for a smallholder to obtain and manage over the certification period. It may also be apparent that evidence for some of the principles such as new plantation development or continuous improvement is difficult to assess, as these require smallholders to produce evidence for attending training programs. Additionally, several principles require field visits and inspection by the auditors,
which can be questioned taking into account the time constraints for such tasks and scope to temporarily satisfy the requirements.

The second key element of RSPO certification process is accreditation to ensure the quality and capability of the auditors. Any organization interested in conducting assessment for certification has to be accredited by the Accreditation Services International (ASI) through a highly stringent, time taking, and costly process. This process though ensures quality and transparency, limits the availability of qualified auditors, which increases the waiting time for willing smallholders and estates. Several large organizations, producers, and, mills have raised their concerns for the lack of sufficient auditors that has become a severe bottleneck in the entire certification process and is significantly limiting the chances of these large organizations to achieve their promised targets by the deadlines. The third key element, as discussed, is the process requirements for RSPO certification that ensures the fulfilment of required criteria and assessment by a qualified and accredited auditor. This requires tracking the documentation and auditors involved in the certification process.

5.1. Limitations of the current RSPO process

Smallholders find it extremely challenging to participate in the prescribed and acceptable manner. Some of the factors instigating this discomfort include absence of information in the right form and medium, lack of understanding of the certification system, age-old farming practices that are difficult to change, inability to quantify the importance of the certification, etc. The key limitations that inhibit the acceptance among the smallholders and scalability of current RSPO certification process are as follows:

5.1.1. Intensive data requirements
RSPO certification requires huge amount of data to be collected and managed at the farm level. This data includes land records, use of pesticides, farming practice, etc. It also requires maintaining records of daily activities on the farms. The smallholders or the farm managers are expected to collect and manage this data in form of various records and proofs. The vast amount of data requirements not only limits the smallholder's capacity but it also limits auditors' capabilities and resources to invest time on the farm activities.

5.1.2. Scope for data alteration (misreporting)
There exist huge motivation and scope for the smallholder for altering or misreporting the data to the auditors to get the certification. There is an upper limit on what can be inspected within the two-day visit by the auditors. For example, a smallholder may be indulged in burning the old-plantations to rapidly clear the farm for new plantation and still may be able to eliminate the traces of burning before the auditor’s visit. Another example may be of employing children on farms, whom the smallholder may avoid employing on the days of the audit. What we are trying to argue is that there are criteria for which the evidence even though provided may not be the representative of the entire activity. More often than not, the auditors have to rely on smallholders’ face-value rather hard-core evidence. In other words, the limitation of current process is to rely on good-behaviour rather hard-core data.

**5.1.3. Change in behaviour**

It requires change in the smallholders’ current farming practice and behaviour, which is highly difficult to achieve. For examples, use of fertilizers and pesticides on the peat land plantations may be behaviours, which may be difficult to change, as smallholders will feel that it will reduce their overall production and thus revenue. Changing smallholders’ daily habits and farming practices is one of the most challenging tasks for the smallholder certification process.

**5.1.4. Lack of qualified auditors**

There is a major constraint on the availability of qualified auditors. The auditors have to undergo a rigorous and costly accreditation process, which limits the number of qualified auditors at any time. Given the huge production of palm oil and demand for certification, there is a severe shortage for auditors. The wait is as long as one to two years. This results in a long waiting time for the smallholders and other producers.

**5.1.5. Lack of transparency**

In the current process there is lack of transparency, as the buyers have to accept that any certified palm oil is sustainable. They have no way to investigate it by themselves. It be large organizations or household buyers, they have no visibility a certain point in the supply chain and have no option but to accept RSPO certified palm oil as sustainable. In some cases, this limits the buyers’ trust and commitment towards the certification, as they might have to pay a premium just
for the certificate because essentially the palm oil is the same. Therefore, it is highly important for the buyers to visualize and be confident that their efforts and premium payments are contributing to the sustainable practices.

5.1.6. Huge cost of certification
There is lack of clarity across the industry about who will endure the cost for this certification. The smallholders assume to receive financial assistance as well as premium for their efforts of getting certified. The buyers assume that the smallholders and palm oil mills shall share the costs, as RSPO will anyway be an industry standard for trade. Consumers assume that the large organizations shall bear the cost from the profits they are earning.

GreenPalm⁴ a certificate trading organization, has introduced a mechanism where the smallholders may be able to sell the certificate independent of the oil to interested buyers. In this case, the cost of certification is paid by the buyers interested in the certificates, while the smallholder continues the palm oil trade as it is. The Greenpalm certificates though encourage sustainable practices are unable to ensure the sustainability claim in the products used by the consumers. These limitations discourage the smallholders as well as the consumers to participate in the RSPO certification process. Data being one of the major input for decision making needs to be better managed. Erroneous data can highly skew the outcomes of the audit. It may have a major impact on the on-farm activities and smallholders’ income from the certificates. In the current process, data entry being conducted by the smallholders on paper-based records is highly prone to entry and interpretation errors. There is high probability of huge amount of documentation errors and personal bias against activates and practices.

There is a need to release the smallholder from the ownership of the data collection and maintenance responsibilities. Then only the smallholder may be willing to participate and will be able to focus on the core farming activities.

There is also a need to encourage the smallholders’ to alter bad-practices by making them aware of the potential good-practices, without directly linking it to the certification process. For example, smallholders’ do not maintain data for the output from each plant, which might help them eliminate non-fruit bearing

⁴ http://greenpalm.org/about-greenpalm
plants and thus reduce the fertilizer intake. Activities like these may be helpful for escalating farm yield and reducing fertilizer consumption, which can satisfy certain criteria in the certification process. But, the smallholder may not be otherwise interested to reduce fertilizer intake just to meet the certification criteria if they fear for a loss in farm yield. Therefore, on one hand there is a need to educate the smallholders about the benefits of the certification system. Simultaneously, on the other hand there is a need to identify potential techniques that can satisfy the certification conditions with minimal alteration of age-old farming practices. The limitations of the current RSPO certification process are presented in Table 2.

<table>
<thead>
<tr>
<th>Main Limitations</th>
<th>Smallholders</th>
<th>Auditors/Buyers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intensive data requirements</td>
<td>Huge time and efforts to collect and maintain data</td>
<td>Not adequate time to review all the data during the audit. Buyers have no visibility of the data</td>
</tr>
<tr>
<td>Scope for data alteration</td>
<td>High motivation to alter or misreport the data</td>
<td>Lack of resources to investigate past events. Buyers have no visibility of on-farm activities</td>
</tr>
<tr>
<td>Change in behaviour</td>
<td>Extremely difficult to change smallholders perception and farming practices</td>
<td>Difficult to evaluate if the behavioural changes are temporary or permanent</td>
</tr>
<tr>
<td>Lack of qualified auditors</td>
<td>High waiting time for audit</td>
<td>Rigorous and costly process that limits the number of auditors</td>
</tr>
<tr>
<td>Lack of transparency</td>
<td>NA</td>
<td>Buyers have no visibility in the overall process.</td>
</tr>
<tr>
<td>Huge cost of certification</td>
<td>Uncertainty about the cost and premium for certification</td>
<td>Buyers not sure about smallholders/ end-consumers willingness to pay</td>
</tr>
</tbody>
</table>

Table 2: Limitations of the current RSPO process

Next we present a big data analytics framework to redesign the certification process.

6. Big Data Analytics Framework

RSPO certification is a highly complex process with intensive requirements for data collection from a wide range of data sources, requiring extra efforts from the smallholders, still being prone to criticism for scope for malpractice. Several of the criteria are almost impossible to inspect in a one day planned visit such as child labour, burning, water level, high conservation value, etc. Critics argue that these can
be temporarily avoided by the smallholders to gain certification but may not be practiced in a long term. Another criticism of the current practice is being retrospective and punitive wherein the smallholders are assessed for the past practices and are penalised for substandard behaviour. They are either certified or not at the end of the audit and yearly reviews, but there is a lack of any mechanism to educate them to prevent these substandard practices. Therefore, though the RSPO is an internationally accepted well reputed certification process and a good starting point, there is a need to re-examine the current certification process and eliminate the scope of malpractices.

There is a need to introduce technology based solutions that can reduce the smallholders' burden for data collection, can reduce the dependency on auditors, reduce the long wait to find an auditor and the results of the audit process, and can make the entire process transparent for the buyers. With the development in sensor and satellite technology, emergence of IoT platforms, and advancement in big data and analytics, it is possible to collect and analyse traditional and non-traditional data. There are a number of existing and proposed data collection solutions in literature and practice. We first identified the data requirements for the current certification process based on the RSPO documents and key informant interview. Then, based on the existing literature, practical applications, and, expert opinion we analysed the alternative mechanisms to collect the data for RSPO certification process. The details are presented in Table 3. It is evident from the table that it is possible to collect the data using advance mechanism. This will eliminate the bottlenecks in the certification process, reduce errors and make it highly transparent even for the buyers. The application of advance technology for data collection across the palm plantations will be generating real time data from a large number of criteria from several hundred thousand smallholders. There will be huge volumes of data generated at a very high rate. This will be highly reliable data of high variety including texts (land records, pesticide use, soil reports, etc.), images (farm images at various point of time), videos (farm activities), voice, etc. The data has all the four characteristics of big data namely volume, variety, velocity, and, veracity. Thus there is a need for big data analytics frameworks that can collect this generated data and process it to conduct descriptive, predictive, and
preventive analysis in real time. Such a system can create accountability, transparency, and, incentive sharing among the supply chain players and enable them to take informed interventions. With the emergence of affordable advance technology, big data analytics has seen enormous uses either it be arts, science, or, commerce. But it is still in its nascent stage for application in agriculture and food production. There exists some stand-alone application using techniques such as remote sensing, network-based investigation, and computer modelling (Jang and Hart, 2015) but there is no supply chain wide application, especially for certification. Therefore, we propose a Big Data Analytics framework to collect and analyse this data for incorporating smallholders in sustainable palm oil certification process. The framework is presented in Figure 3.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Evidences</th>
<th>RFID</th>
<th>Satellite, and UAV Sampling</th>
<th>GPS and Soil Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use of Fertilizers</td>
<td>Records of fertilizer input shall be maintained</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Use of Pesticides</td>
<td>Records of pesticide use</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Store pesticides according to best practices</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Document all pesticide aerial application</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Justification for use of all pesticides</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Use proven methods of pesticide application</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pesticides to be used by trained personnel only</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Environmental Concerns</td>
<td>Evidence of periodic tissue and soil sampling</td>
<td></td>
<td></td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>Protection of water courses and wetlands</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HCV assessment on planted area</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Take measures to maintain HCVs and RTEs</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td></td>
<td>On going monitoring plan for program</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Waste Management</td>
<td>Identify and document all waste products</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dispose all chemicals</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
</tbody>
</table>
There are several researchers who have proposed big data analytics framework based on the applications (Tekiner and Keane, 2013). As discussed by Miller and Mork (2013) the framework can be divided into three key stages namely: data discovery, data integration, and, data exploitation. Data discovery involves, data collection, preparation, and organization. At this stage, data from multiple sources is collection, thus there is a need to filter, integrate, and process the data. As discussed earlier, there are more than 140 criteria for which evidence are required for the certification process. These evidence or data points include a wide-array of activities which will generate structured as well as unstructured data that will include not just text but also images, voice, and videos. Some of the data such as smallholders identification, land records, plantation age, etc. may be obtained in a structured data cubes or arrays. Most of the other data from the day-to-day activity have to be collected and mined to extract the required data points. Given the current state of records, satellite image processing capabilities, sensor technology, it is highly likely that there might be a need to condition and pre-process the data to eliminate any potential error.

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Procedural disposal of waste material</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Waste management plan to be in place</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Fossil fuel efficiency use plan to be in place</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Deforestation Concerns</td>
<td>No burning of plan except under ASEAN guidelines</td>
<td>Y</td>
</tr>
<tr>
<td>Where fire is used to prepare land, there shall be evidence of approval</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Conduct assessment of all polluting activities</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Identify GHG emissions and mitigate them</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Monitoring system should be in place</td>
<td>Y</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Use of advanced technology for data collection
The second stage involves data integration into uniform representation to enable meaningful analysis. Integrating the data extracted from multiple sources into a uniform database will be a major roadblock. We propose to use a Master database (metadata repository) to integrate the most useful data for current application and future reference. The master database facilitates discovery of hidden patterns and trends. These patterns and trends can help identify the structural difficulties that are otherwise invisible. There are techniques such as data federation and semantic web technologies that enable data integration (Miller and Mork, 2013).

The data exploitation stage involves analysis, visualizations, and, decision support. The collected data such as smallholders’ daily farm activities, adaptation of sustainable practices, farm yields, etc. will be very unique given the inherit uniqueness of the smallholders from which the data will be collected. It is also to be noted that this data will be highly sensitive as it contains a very detailed profile of each smallholder in the certification system. Thus there is a need to collect and analyse the data in a way that can help individual smallholders while maintaining their privacy. There is also a need to establish data protection system and access-levels for individuals. Once the data is ready for analysis, descriptive analysis shall be conducted that will include data visualised, reporting, and, sorting to identify the obvious trends and outliers. Based on the requirements for the certification process, data can be analysed to identify the hidden patterns and dynamics. This step will address the requirements for the certification process, as the reports can be used to determine the credentials for an individual smallholder for RSPO certification.
Figure 3: Proposed Big-Data Analytics Framework for Sustainable Palm Oil Production
7. Discussion

This section discusses the requirement and relevance of the proposed big data analytics framework. There are organizations such as Unilever and Mars that on their own are committed to move beyond RSPO. There is also significant pressure from NGOs such as Greenpeace to look beyond RSPO. As discussed earlier, the existing RSPO certification process despite being a good first step has severe limitations such as huge costs and limited auditors that constrain the speed of certification. One of the most affected groups were the smallholders who find it extremely difficult to arrange for funding and auditors to get their plantations certified. In absence of resources they become highly vulnerable and are generally ousted from the sustainable procurement efforts of the large organizations. Several organizations such as P&G, Colgate, Nestle, Johnson & Johnson, etc. have committed to procurement of sustainable palm oil and to achieve 100% deforestation in their palm oil supply chains by 2020. There are other organizations such as Pepsi, Godrej, General Mills, which are currently under tremendous NGO and competitor pressure to commit to procurement of 100% sustainable palm oil. It is expected that soon there will be many more companies committing to procurement of 100% sustainable palm oil making it an industry standard. This will further fuel the demand for certified palm oil creating pressure on the existing limited certification resources. Technically, the certification verifies the process to be sustainable while the product (palm oil) remains the same. Thus it is highly essential to maintain transparency and traceability in the certification process.

Moreover, as the auditors are paid by either the company or the farmer thus in order to maintain consumers’ trust there is a need for an independent source to verify the claims made by the companies.

The proposed big data framework seems a potential solution to address the current limitations and concerns for transparency. Implementation of the framework will reduce the dependence for the auditors to visit the farms for physical inspection rather they will be able to analyse the collected data to make a decision for the smallholders’ application. This will surely enhance the speed of certification, but the process still remains descriptive in nature where it will be reporting what has already happened. This will develop anxiety and uncertainty among the smallholders and buyer about the outcome of the periodic review. There is also a need to eliminate or reduce the level of uncertainty. The proposed big data framework can conduct predictive
analysis on the pre-processed data to predict the future trends and potential smallholder responses to these trends. This will be highly useful for the buyer as they can identify the smallholders at an early stage to determine the potential number of smallholders that might get RSPO certified.

Both the descriptive and predictive analysis are punitive in nature. In both of these cases, if a smallholder adopts any unsustainable practice or fails to adhere to the stringent RSPO standards, S/he will be eliminated from the process either early or at the end. They will not get the certification and in a way will be labelled unsustainable. This will demotivate the smallholder to participate in the process and will force them to sell their produce in the open market. But the objective is not just to identify smallholders who can provide sustainable palm oil but is to eliminate unsustainable palm oil practices. Therefore, exists a need to educate the smallholder at the right time about any potential violation due to his action. For example, if a smallholder is setting up the farm on fire to clear the old plantation, there is a need to inform him and if possible stop him at the moment otherwise it will not only disqualify him but also will be a damage to the environment. The proposed framework can assist by processing real-time data and generating preventive actions for the stakeholders. For example, if a smallholder is burning the farm, data from UAV or sensors can be collection about

Figure 4: The application of the proposed big data analytics framework
the smoke or temperature and can be quickly processed and an alert can be send to the smallholder and farm manager (local youth). This will not only abstain the smallholder from accidental violation but also reduce the chance for disqualification from the certification process. Details of the three applications namely descriptive, predictive and preventive are presented in Figure 4.

8. Conclusion and Scope for future research

This research critically analysed the current RSPO certification process for smallholders to ensure sustainable palm oil production. It was found that there are severe limitation of the current process in terms of the capacity constraints and resource requirements. Moreover, the process lacks transparency, accountability, and, traceability. There was a need to overcome the limitation to enhance smallholder participation and customer willingness. Thus we analysed the current process and data sources required for certification. Based on the findings, potential alternate means of data sources were identified and a big data analytics framework was proposed. It is expected that this framework will be highly useful for the producers, buyers, policy makers, and, RSPO. The framework will convert the current punitive approach to a more preventive approach. There are certain inhibitors of the proposed framework and needs to be addressed for it to be a viable solution. This research makes a novel and valuable contribution to the literature by critically evaluating the RSPO certification process to highlight the complexity, difficulties, and, uncertainties that constraint the smallholder's participation. There exists very limited literature that addresses the 'sustainability certification schemes' from a business process perspective. It also highlights the limitations of these schemes in terms of the ambiguity of certification, and sustainability assurance to the end consumers. There exists papers that propose certifications as a final solution, but very few have evaluated the certification process itself. There also exists literature that measures the confidence of end consumers on certification schemes but very few have studied the reasons for low confidence. This paper makes a unique contribution by considering a data driven business process analysis approach to identify the issues in incorporating the smallholders in sustainability initiatives. By critically analysing the certification
process, data requirements, and data ownership, this research was able to propose a big-data analytical framework to incorporate the smallholders in sustainable production. There exists no literature that proposes the use of big-data analytics to address the issues in sustainability certifications. The proposed solution is highly significant as it not only address the existing concerns for the smallholder and buyers but also provide more visibility and transparency to the end consumers, NGOs, and policy makers.

The proposed big data analytics framework presents tremendous implications for the managers, customers, and policy makers. The managers can use the framework to develop their plan to meet their targets for sustainable palm oil procurement. They can also work closely with the suppliers and have higher control and visibility of the entire process. They can use the framework to eliminate the risk of having unsustainable palm oil in their system. They will be able to better align the incentive system to encourage the suppliers. The framework will provide higher transparency and traceability to the consumers, thus will increase their willingness and commitment towards sustainable palm oil. The framework will facilitate the policy makers to device interventions to further enhance the sustainability practices across the farms. The framework will also be able to provide evidence for the benefits of RSPO implementation. RSPO member can evaluate the effectiveness of the proposed framework and its ability to scale RSPO’s efforts.

It is expected that the proposed framework will help the smallholders get not only the certification but also premium for their efforts. The framework will reduce the data burden from the smallholders and will help them enhance the farm practices while focusing only on the farm activates. Its preventive nature will reduce the elimination of the smallholders from the process thus will have a significant positive impact on the environment which otherwise was not possible.

The proposed big data analytics framework, though has the potential to address the current key concerns, has few challenges for its implementation. As the data collection will require application of advance technology, there will be a need for initial investment. This may be one of the biggest challenges to gain financial support due to the lack of clear ownership. There are few successful applications such as forest watch that was developed by World Resource Institute (WRI) and uses satellite images and google map to collect data for further analysis. Another major challenge
may be to define a governance structure for data ownership. All the stakeholders would like to have access to all possible data but that may risk the interest and identity of others. Additionally, there may be few data points for which it might be difficult to attain acceptable volume or quality of data using existing technology. The third major challenge is the lack of skills to design, develop, and deploy the proposed framework in an efficient manner. There may be other concerns from the stakeholder who might be directly affected by the framework such as the auditors.

In future, researcher may wish to explore the economic viability of the proposed big data framework. They may also explore the changes in buyer-supplier relationship post certification

**Acknowledgement:** The authors sincerely thank the anonymous reviewer(s) for their critical and constructive comments that significantly helped to improve this paper.

**Reference:**


15. Everaerts, J., 2008. “The use of unmanned aerial vehicles (UAVs) for remote sensing and mapping”. *The International Archives of the*
*Sustainable Development, 22(5), pp.289-310.*


