Oil in the Sahara: mapping anthropogenic threats to Saharan biodiversity from space

Clare Duncan1, Daniela Kretz1,2, Martin Wegmann3,4, Thomas Rabeil5 and Nathalie Pettorelli1

1Institute of Zoology, Zoological Society of London, Regent’s Park, London NW1 4RY, UK
2BayCEER, University of Bayreuth, Bayreuth 95440, Germany
3Department of Remote Sensing, University of Würzburg, Campus Hubland Nord 86, Würzburg 97074, Germany
4German Remote Sensing Data Center, German Aerospace Center, Oberpfaffenhofen 82234, Germany
5Sahara Conservation Fund, Niamey, Republic of Niger

Deserts are among the most poorly monitored and understood biomes in the world, with evidence suggesting that their biodiversity is declining fast. Oil exploration and exploitation can constitute an important threat to fragmented and remnant desert biodiversity, yet little is known about where and how intensively such developments are taking place. This lack of information hinders local efforts to adequately buffer and protect desert wildlife against encroachment from anthropogenic activity. Here, we investigate the use of freely available satellite imagery for the detection of features associated with oil exploration in the African Sahelo-Saharan region. We demonstrate how texture analyses combined with Landsat data can be employed to detect ground-validated exploration sites in Algeria and Niger. Our results show that site detection via supervised image classification and prediction is generally accurate. One surprising outcome of our analyses is the relatively high level of site omission errors in Niger (43%), which appears to be due to non-detection of potentially small-scale, temporary exploration activity: we believe the repeated implementation of our framework could reduce the severity of potential methodological limitations. Overall, our study provides a methodological basis for the mapping of anthropogenic threats associated with oil exploitation that can be conducted across desert regions.

1. Introduction

The human population and its demand on the Earth’s natural resources have increased exponentially over the past century, leading to a global biodiversity crisis [1–4]. With present levels of biological diversity drastically altered from their historic states [1] and declines predicted to continue under growing pressures from land-use change and climate change [2,5,6], scientists and conservationists are now grappling with the immense challenge of monitoring and halting the current rate of global loss [7]. A number of prominent articles have made the case for targeting conservation funding at biodiversity hotspots, tending to fall predominantly within tropical biomes. The reasoning behind such argumentation is that management actions there target the greatest number of species per surface unit, thus maximizing cost to benefit ratios [8–13]. This focus on hotspots has inevitably resulted in the neglect of less species-rich biomes [14–16] where threats are poorly documented and monitored, and their local, regional and global impacts on biodiversity are poorly known, leading to a much reduced understanding of the potential need for conservation action in these regions [14–16].

Deserts are one such biome that has previously been poorly studied and has received little attention from conservation activities [14,15], despite covering 17% of global land mass and harbouring surprisingly high levels of biological diversity, including many endemic species and some of the most endangered species in the world [17–20]. This neglect is worrying as (i) available...
information suggests that desert and dryland biodiversity is declining rapidly [14,18–22]; and (ii) predictions show that the rate of climate change is likely to be particularly high in the desert biome, with species potentially facing especially difficult adaptation challenges in the near future [23]. This is of particular concern, as the ecology and distribution of desert-adapted species are under extreme climatic control [24]. Other major threats to deserts include overgrazing, woody-vegetation clearance, agricultural expansion, water diversion and extraction, soil and water pollution, land conversion due to industrial activities and associated threats from armed conflicts [19,21]. With potentially enormous local- and regional-scale effects, a current major anthropogenic threat to desert biodiversity, likely to increase in extent and intensity over the coming years through increasing resource demand, is oil exploration activities [21,25,26]. Across the Sahara and Sahel, threat from oil exploration is particularly great, with concerns for global oil stocks having led to exploration being undertaken in increasingly remote areas [27,28].

The future of Sahelo-Saharan biodiversity is considered to be largely dependent on anthropogenic expansion in the region [21]. Countries like Niger, which represents a stronghold for local desert species [21,29,30], have undergone rapid increases in their oil extraction and refinery activity in recent years [26,30]. Such increased oil prospection activities have, for example, now led to the endangerment of the last known viable population of addax (Addax nasomaculatus), owing to habitat loss in important, high-quality grazing sites in the Tin Toumma desert [31]. These largely unmonitored activities are expected to occur across many parts of the region, with potential ancillary impacts including waste production, reduced water availability through anthropogenic extraction, reduced habitat quality and ecosystem functioning, and increased poaching and human–wildlife conflict [21,30,32,33]. These potential threats could have an enormous impact on local biodiversity, as (i) much of the fauna in the Sahelo-Saharan region is considered to be of global conservation concern (e.g. the addax, Dorcas gazelle (Gazella dorcas), dama gazelle (Nanger dama), Saharan cheetah (Acinonyx jubatus ssp. hecki) and Kleinmann’s tortoise (Testudo kleinmanni); [18]) and (ii) the distribution of biological diversity in the area is markedly fragmented, with often highly localized hotspots [20,21] that could become imperilled by the impacts of unmonitored oil exploration activity. Hotspots of Sahelo-Saharan biodiversity are often concentrated within protected areas (PAs) [21]. If conducted close to or within these areas, the ancillary threats and associated impacts of oil exploration activities could lead to drastic reductions in wildlife populations, reducing the effectiveness of PAs in conserving these remnant hotspots of biodiversity. Across the African continent, the designation of oil concessions within PAs is common, and particularly within those with a high level of protection [34]. Threats from oil exploration activities at PA boundaries could also result in reduced connectivity of habitats between these PAs and important resource areas by reducing the quality of habitat corridors and producing physical barriers to movement, cutting off wildlife populations from wider ecosystems and satellite populations [35–38]. Such impacts on wide-ranging animal species could have potential knock-on cascading effects through other trophic levels within such systems [39,40], producing broad-scale ecological impacts and impacts to overall ecosystem health and functioning. Despite these potential consequences, little is currently known about how and where oil exploration and exploitation activities are taking place, and where these developments are likely to negatively impact Sahelo-Saharan ecosystems.

A major issue in determining threats to biodiversity from the expansion of oil exploration activities is the difficulty in obtaining field data on the geographical locations of these developments. Field monitoring can be notoriously costly, and field data on oil exploration and exploitation activities, while accurate, are difficult to obtain for large areas, especially for remote regions where logistics and safety are real issues [41,42]. In light of this, remote sensing technologies could provide a useful monitoring alternative. Global-scale low-light imaging satellite data from the US Air Force Defense Meteorological Satellite Program Operational Linescan System [43], data from the European Space Agency’s Along Track Scanning Radiometer sensors [44] and data from the recently launched VIIRS instrument on the Suomi National Polar Partnership satellite [45], can be used to derive global and national information on oil exploration activities. However, the spatial resolution of these datasets is relatively coarse (0.56 km, 1 km and 750 m, respectively); thus, information derived from these data may not be applicable to identify fine-scale, local-level oil exploration threats to biodiversity. Very high-resolution satellite imagery could instead be used to derive national information on oil exploration and exploitation activities; yet, most organisations and institutions in need of such data are unable to afford the cost of data acquisition and processing at such spatial extents. Freely available satellite data such as data from the National Aeronautics and Space Administration (NASA) and United States Geological Survey’s Landsat satellites, however, have shown promising applications in the fields of ecology and conservation [46] and could potentially inform efforts to detect and map oil exploration and exploitation activities at relevant scales [47,48]. Furthermore, the availability of Landsat satellite data at a good and consistent temporal scale (at up to 16-day intervals) can allow for the continued and repeated monitoring of areas of concern.

Surprisingly, the potential for freely available satellite data to inform conservation efforts in desert systems currently remains largely unexplored (but see [47]). This study aims to fill this knowledge gap, by assessing if and how open access satellite data can be reliably used to monitor oil exploration and refinery activity in the Sahelo-Sahara region. It aims to do so in a reproducible manner, so that methods and analyses developed herein can be applied to conservation activities in other dryland regions. Many approaches and attempts to detect the location and distribution of oil exploration and extraction activities focus on the detection of oil flares alone [43–45]. However, sites also feature buildings and other man-made structures, which may include water ponds, roadways or visible smoke [46], and these features are associated with different spectral signatures than those generally found in the surrounding desert [47,49]. As such, a framework that aims to detect not only oil flares but also associated structures may improve the success of detection of novel and established oil extraction activity. This serves as the basis for our approach, which aims to validate the use of Landsat satellite data for the detection of features typically associated with oil exploration activities. Our study undertakes a two-step approach: firstly, field-derived (known oil exploration site locations) and visually derived (high-resolution Google Earth imagery) validation data, and satellite-based landcover classification are combined to establish a
framework for the detection of known features associated with oil exploration activities; this framework is then applied in areas for which an explicit landcover classification cannot be made, in order to provide spatially explicit information about recent development in these activities in relation to existing PAs.

2. Material and methods

(a) Study areas

The location used for the development of our monitoring framework is in Saharan Algeria. The area of interest reaches from the northwestern corner at 31°49'12" N, 8°22'12" E to the southeastern corner at 31°4'12" N, 9°27'0" E (figure 1). A second study area in Niger, where our landcover classification is then implemented, is delimited by a box reaching from the northwestern corner at 16°22'12" N, 12°12'0" E to the southeastern corner at 15°27'36" N, 13°49'12" E (figure 1). Both study areas comprise areas of desert, rocky terrain and sparse vegetation. Although they both lie within the Sahara, Algeria and Niger represent two extremes in a spectrum of anthropogenic impacts in the hyper-arid zone: Algeria has a great number of established oil and gas extraction sites, as well as refineries [21], and here represents areas with high anthropogenic pressure associated with oil exploration activities. On the other hand, desert habitat in Niger remains largely undisturbed by anthropogenic activities, and the expansion of oil exploration and exploitation activities in the country has occurred only in recent years [30,50]. The Niger region was selected for this study as it borders key PAs for the conservation of desert biodiversity; namely, the Air & Ténéré National Nature Reserve, and the Termit & Tin Toumma National Nature Reserve (TTTNNR), which are currently threatened by the expansion of oil exploration activities [30,50]. The study area in Niger encompasses part of the eastern side of the TTTNNR and also the Agadem oil concession block, which is the predominant site of emerging exploration activity in the region to the east of TTTNNR (figure 1; [30]). Within TTTNNR and Air & Ténéré, much of the remaining Sahelo-Saharan diversity thrives [18,30,50], including 18 large mammal species, 19 reptile species, 137 resident and several Afro-tropical and Palaearctic migrant bird species [51]. These two PAs protect important populations of many species of global conservation concern: carnivores (the Saharan cheetah (population estimate currently less than 10 individuals in the area; [51]), striped hyaena (*Hyaena hyaena*) and the greatest diversity of small and medium sympatric carnivores in the Sahara; [51]), birds (the Nubian bustard (*Neotis nuba*) and Arabian bustard (*Ardeotis abu*)) and large ungulates (the addax, Dorcas gazelle, dama gazelle and Barbary sheep (*Ammotragus lervia*);
[18,30,50]). Importantly, the Termit massif and Tin Toumma desert region hold the last remaining viable population of addax in the world (ca 200 individuals; [52]). The direct ecosystem impacts of increased oil exploration activities [21,30,32,33] have the potential to produce profound effects on the local biodiversity; habitat loss and water extraction at prospect sites may serve to affect floral abundance and composition, in turn reducing habitat quality and resource availability for species at higher trophic levels [21]. Moreover, increased poaching of large ungulate species [30] may serve to reduce their densities in such areas, and in turn, coupled with increased human–wildlife conflict, may result in the reduction of populations of key Saharan carnivore species, e.g. the Saharan cheetah, where ungulate prey are a limiting factor to predator abundance [53]. In addition to the ecological uniqueness of the Termit massif and Tin Toumma desert area, it is also thought to be rich in currently largely unstudied archaeological, anthropological and palaeontological history [30,50].

Despite the discovery of hydrocarbons within Niger in 1975, little oil exploration and exploitation activity has previously occurred in the country, due to the low ratio barrel price (T. Rabeil 2013, personal communication) and a lack of economic incentive caused by a historic wealth of uranium providing the country’s main economic export [30,50]. However, concerns over diminishing uranium stocks over the past two decades have led to the designation of oil concession blocks throughout much of the country [26,30,50]. In June 2008, the government of Niger signed a contract with the Chinese National Petroleum Corporation (CNPC) for the purchase of three oil concession blocks covering large areas of the Air & Ténéré and the TTINNR (figure 1; [30,50]), with oil exploration activities occurring within these areas from June 2008 onwards. With extremely little ecotourism in the country, the societies and government of Niger have little economic incentive to conserve the ecological, archaeological and anthropological uniqueness of the Termit massif and Tin Toumma desert region [50]. At present, 20 oil concession blocks have been designated across the country, with four currently under licence (including the three rented by CNPC and one by the Algerian corporation SONATRACH) covering much of the Termit massif and Tin Toumma desert region (figure 1; [30]). Although most designated concession blocks in Niger remain open at present, depleting uranium stocks are likely to result in an economic shift to an increased reliance on hydrocarbon concession rentals and exports [30,50], and thus increased sale of concessions to oil corporations, extending the current level of threat to local biodiversity. The expansion of these activities into the neighbouring PAs is of particular concern for large ungulate species as (i) many of these are highly threatened remnant populations and (ii) there have been reports of poaching (gazelles and birds) associated with the arrival of workers for oil exploration activities in the region since 2008 [54]. Ancillary impacts associated with industrial soil and water pollution have also been found to be evident at exploration sites within the oil concession blocks [55].

### (b) Satellite data

Our methodology makes use of Landsat 7 Enhanced Thematic Mapper Plus (ETM+) imagery, available at 30 m spatial resolution ([56]; available for download at: http://earthexplorer.usgs.gov), which has previously been found to have utility in desert areas [47,57]. One minor limitation of the Landsat 7 ETM+ sensor is that the Scan-Line-Corrector is off (SLC-off) owing to a technical malfunction, causing approximately 22% of data from any given Landsat 7 ETM+ image to be lost [58]; ‘striping’ of missing data is apparent in the images. Typically Landsat 5 TM imagery would be applied instead due to this error; however, owing to limited data availability over Niger, we were unable to obtain Landsat 5 TM imagery for analysis here, and thus Landsat 7 ETM+ data were used for both regions despite the SLC-off issue. The Landsat 7 ETM+ scenes considered for both study areas were affected by the SLC-off issue; however, gap masks facilitated the masking out of the affected region, and thus the SLC-off issue can be resolved for use in landcover classification [56]. The images obtained for these analyses were from 5 November 2012 for the Algerian area (framework development area), and 2 November 2012 for Niger (figure 1), in order to avoid differences according to varying seasonal solar elevation and soil–water–vegetation dynamics [59].

### (c) Field-derived and visually derived training and validation data

In order to produce and validate a landcover classification from satellite imagery, field-derived and visually derived validation points of specific landcover classes can be used. In both Algeria and Niger, ground-truthed geo-referenced (GPS-logged) site validation point data were obtained from the Sahara Conservation Fund (SCF) over the period 2009–2011, detailing the locations of known oil exploration and extraction activity. These data included one permanent site within the Algerian Landsat 7 ETM+ scene. In addition, we located two further validation sites in the Algerian study area through inspection of very high-resolution QuickBird and IKONOS imagery within Google Earth Pro [60] as performed in the work of Fritz et al. [61], resulting in a total of three validated oil exploration sites in this study area (table 1). Using the very high-resolution Google Earth imagery, we were also able to identify a total of eight landcover types in the scene: desert, bare rock, vegetation, roads, dark-coloured settlements, light-coloured settlements, oil flares and water ponds. These eight landcover classes were selected based on the study aims of differentiating human infrastructure from natural landcover (i.e. bare rock, vegetation and desert versus roads, settlements, oil flares and water ponds), in order to develop an appropriate landcover classification. An equal distribution of training point data samples were then created across all landcover classes to define their relative spectral information from the Landsat 7 ETM+ scene (200 random samples of pixels per class) and were then used to develop a landcover classification model.

<table>
<thead>
<tr>
<th>site locations</th>
<th>Algeria</th>
<th>Niger</th>
</tr>
</thead>
<tbody>
<tr>
<td>field derived (SCF)</td>
<td>1</td>
<td>7*</td>
</tr>
<tr>
<td>visually derived (Google Earth)</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

*One of the original eight field-derived sites in Niger was no longer present in November 2012 (a temporary site), while the status of the remaining seven sites is not known from SCF; thus, these sites were considered to be ‘permanent’.

5 TM imagery for analysis here, and thus Landsat 7 ETM+ data were used for both regions despite the SLC-off issue. The Landsat 7 ETM+ scenes considered for both study areas were affected by the SLC-off issue; however, gap masks facilitated the masking out of the affected region, and thus the SLC-off issue can be resolved for use in landcover classification [56]. The images obtained for these analyses were from 5 November 2012 for the Algerian area (framework development area), and 2 November 2012 for Niger (figure 1), in order to avoid differences according to varying seasonal solar elevation and soil–water–vegetation dynamics [59].
from SCF beyond 2011, and these are thus considered here as ‘permanent’ extraction sites. Recent very high-resolution Google Earth imagery was not available for the Niger study area, and thus associated landcover classes could not be visually identified and classified for this area, leading to the prediction of landcover here by using the classification model developed for the Algerian study area scene.

(d) Statistical analyses
All analyses were performed in GRASS v. 6.4.2 [62], QGIS v. 1.8.0 [63] and R v. 3.0.0 [64], using the packages ‘raster’ [65], ‘rgdal’ [66], ‘randomForest’ [67] and ‘spgrassix’ [68], the latter of which serves as an interface between R and GRASS. A conversion of digital numbers to top-of-atmosphere reflectance values was carried out on the raw Landsat 7 ETM+ scenes. Owing to a lack of strong elevational gradients across both study areas (Algeria: 100–300 m a.s.l.; Niger: 300–400 m a.s.l.), no further orthorectification of the images was necessary [69]. Texture analyses were conducted on the Algerian Landsat 7 ETM+ scene using a moving window approach, in order to emphasize features typical of oil refineries and highlight them against the background of the Saharan desert [70–72]. A moving window size of 5 × 5 pixels has been found to be useful in highlighting changes in edges of landcover classes in arid environments [73]. Such a window size (approx. 150 × 150 m) would also encompass the area of a small oil refinery but is not so large as to neglect roads; this size was thus considered in these analyses. The texture analysis included the mean, maximum, minimum, variance and standard deviation of all eight Landsat 7 ETM+ bands (band 6 is composed of two separate bands). All metrics are based on the neighbourhood statistics within the ‘r.neighbours’ module of GRASS (with square orientation). Thus in total, 40 parameters (eight bands × five moving window variables each) were calculated. These were used as input data for the supervised landcover classification of the acquired scenes [41]. Landcover classification was developed using a Random Forest model approach [74], based on the model training data (200 pixels for each landcover class) mentioned above, using the aforementioned calculated parameters (40 in total) as the independent variables for Algeria [75].

The relative importance of model variables within the Random Forest model (i.e. the contribution of each variable in the model in the construction of each decision tree) was assessed for each variable by examining the mean decrease Gini index (MDG; [76]); variables with higher MDG have greater importance in a Random Forest model. In addition to the Random Forest standard construction of 500 decision trees to determine the best model for landcover classification, we also aimed to assess the stability of the variable importance across models. To do so, we ran the Random Forest model 200 times and observed changes in the variable importance parameter of the MDG over the runs. Random Forest model accuracy was assessed by examining the model internal class error rates (defined as the percentage of the 200 pixels falsely classified) for each landcover class identified within the model, as well as the ‘out-of-bag’ estimate for the generalization error for the overall model [74]. The Random Forest classification model was then further validated by comparing model-detected oil exploration sites with the previously mentioned validation data from ground-truthed geo-referenced data points. The Random Forest model derived and validated on the scene acquired for Algeria was then applied to the Landsat 7 ETM+ scene for Niger [67]. Validation of the Niger classification used a combination of the aforementioned a priori-selected ground-truthed validation sites and, in the absence of recent very high-resolution Google Earth imagery for the area, a visual red–green–blue (RGB; bands 7, 4 and 2) image assessment of the original Landsat 7 ETM+ data for the considered area.

3. Results
The Random Forest model accurately predicted the distribution of all eight landcover classes for Algeria. Model internal class errors ranged between 0 and 0.02%: bare rock, oil flares and water all had 0.00% error; vegetation, roads, and dark- and light-coloured settlements all had 0.01%; and finally, desert had the highest error rate at 0.02%. Moreover, the out-of-bag overall model error rate was 0.62%. On average, the Landsat ETM+ bands providing the most important variables were band 3 (red), followed by band 1 (blue), band 7 (short-wave infrared-2 (SWIR-2)), band 4 (near-infrared), band 6–2 (thermal-infrared 2), band 6–1 (thermal-infrared 1), band 2 (green) and band 5 (mid-infrared; figure 2; electronic supplementary material, Appendix S1). The field-derived refinery location in Algeria was accurately predicted by the Random Forest model and visually confirmed using high-resolution Google Earth imagery, and the two sites located via high-resolution Google Earth imagery were also accurately predicted by the Random Forest model (table 2). Detecting the accurate location of these sites was made possible by the presence of concentrated areas of non-desert landcover classes: oil flares, settlements, roads and water ponds.

When applying the identification framework to the scene from Niger, a total of seven oil exploration sites were identified (figure 3 and table 2). Of the detected sites, four oil refineries were field validated by existing geo-referenced field data from SCF; thus, three new oil exploration sites were discovered from our analyses (figure 3 and table 2). RGB visualization (7, 4 and 2 bands) of the Landsat 7 ETM+ image, confirmed that the newly detected sites were consistent with other validated sites. However, some of the ground-truthed geo-referenced exploration sites from SCF in Niger were not detected when applying the Random Forest model framework (omission error 43%; figure 3 and table 2). In contrast to oil exploration sites in the Algeria study areas, aside from desert there were no other discernible classes of landcover located near the oil flares identified in Niger. The oil flares class thus enabled the detection of the oil exploration and exploitation sites in the Niger region, which is due to the unique spectral and thermal signal of the class; surface temperature (K, calculated from band 6–1 of the Landsat ETM+ sensor [77]; figure 4) was highest for this landcover class, and high IR values (maximum band 7 (SWIR-2)), which contributed greatly to the Random Forest classification (figure 2; electronic supplementary material, Appendix S1), are associated with the heat signature of the flame [47,57].

Ground-truthed SCF point data for the Niger region show an existing overlap between PA boundaries and oil exploration sites (within TTTNNR; figure 3), however, the Random Forest identification model applied here did not detect any activity within the part of the PA that fell within our study area and did not detect the two oil exploration sites within the study area–PA overlap (10% of the study area overlapped with the PA; figure 3). The proximity of many of the model-detected oil refineries to the nearest boundary of the TTTNNR was found to be less than 15 km (figure 3).

4. Discussion
Owing to limited funding and conservation attention [14–16], little is currently known about the location and intensity of
specific threats to biodiversity in desert ecosystems. Despite being considered a major threat [21], and increasing licensing of concessions [26–30,50], especially in close proximity to PAs (figure 1; [21,30,34,40]), at present no monitoring framework exists for the expansion of oil exploration activities in desert regions. As a result, there is presently a lack of realtime information regarding where conservation efforts should be targeted in order to determine the potential impacts of these activities on local biodiversity. The results of this study provide a framework to assist in closing this knowledge gap, showing for the first time that current and expanding oil exploration activities in desert ecosystems can be detected and mapped remotely via freely available satellite data. Our results confirm that predictions from ground-verified land-cover classification models (here Algeria) can be successfully applied to detect activities in areas for which classification models cannot be explicitly produced (here Niger). The methodology developed herein thus makes the introduced framework reproducible across other desert regions and in areas for which location-specific classification models cannot be constructed owing to limited availability of high-resolution imagery and ground-truthed data for validation. This is a situation that is particularly common in many remote desert regions. Owing to the free availability of NASA’s Landsat satellite data, the framework developed in this study is inexpensive and easily implementable across deserts, and thus may be used to monitor oil exploration activities over time in a cost-effective manner in areas for which conservation funding is limited [15,16]. Moreover, the spatial resolution of the input variables (150 m, based on the original 30 m Landsat 7 ETM+ bands) is much finer than other datasets traditionally employed for the detection of oil flares [43–45], and thus is more appropriate for identifying anthropogenic structures, while omitting small-scale landscape variations (e.g. singular rocks), in order to determine regional-scale expansion in the distribution and intensity of activity. Detecting previously unknown oil exploration sites in close proximity to PAs (figure 3), the framework developed herein enables the mapping and continuous monitoring of novel and intensifying threats to localized hotspots of desert biodiversity.

The successful application of the model classification and identification framework developed in this study can be attributed to the 40 input variables derived from the eight Landsat

---

**Figure 2.** Random Forest model variable importance; the MDG. All 40 variables are sorted according to their median importance in the model after 200 runs. The most important variables in the Random Forest model have a median MDG between 30 and 40. Variables are depicted with decreasing importance from top to bottom.

**Table 2.** Random Forest model classification results for the model training study site in Algeria and when applied to the additional study site scene in Niger. Omission errors for model detection of oil exploration and extraction sites from field-derived locations. Additional ‘novel’ model-detected sites in Niger that were not known from original field-derived location data are also listed; these were visually validated using an RGB visualization (7, 4 and 2 bands) of the original Landsat 7 ETM+ scene.

<table>
<thead>
<tr>
<th>area</th>
<th>field-detected sites</th>
<th>model-detected sites</th>
<th>omission error (%)</th>
<th>novel sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algeria</td>
<td>1 (2)</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Niger</td>
<td>7</td>
<td>4</td>
<td>43</td>
<td>3</td>
</tr>
</tbody>
</table>

*aAdditional sites in brackets determined via high-resolution Google Earth imagery.*
7 ETM$^+$ spectral bands. Class separation was supported by the moving window variables, as was found by Ge et al. [41]. Overall, texture analyses have been found to enhance information from the raw satellite bands [71]; the lightness or darkness of certain objects in relation to neighbouring objects can be identified for the different bands. In using texture analyses, the importance of variables is dependent upon the targeted object of the analyses [71,78]. However, the most important variables for this framework (based on the MDG; figure 2; electronic supplementary material, Appendix S1) seem to be the band averages, which indicates that the bands themselves are still highly important in predicting landcover classification. We do, however, acknowledge that taking the average values also reduces the effective spatial resolution of the input variables, producing a trade-off between the detection of large-scale features versus increasing classification error according to mixed pixel signals. The minimum band values were also shown to play an important role in image classification, as landcover classes such as water, vegetation, dark buildings and road tarmac typically have lower spectral values in comparison to other classes [59]. While in Algeria, identification of oil exploration and extraction sites was made possible by the presence of assemblages of multiple landcover types (oil flares, small water ponds, settlements and roads), within the Niger study area such associated landcover classes were not present and sites were made detectable by the presence of oil flares. Oil flares have indeed previously been found to be good indicators of refineries in the Sahara [33]; here, these could be visualized both in the raw satellite images (as in the Kuwait oil fires; [47]) and with classification results. Such detection is primarily due to the band 7 (SWIR-2) of Landsat 7 ETM$^+$ sensor, which covers a similar spectral range as the thermal IR Advanced Very High-Resolution Radiometer band that has been previously found to be successful in detecting oil fires [49]. The maximum value of band 7 (SWIR-2) was particularly useful in the model (figure 2; electronic supplementary material, Appendix S1), which provides information on the heat signature of the oil flare [47,57]. The successful application of this identification framework is also largely a result of the chosen study region; in desert ecosystems, cloud contamination of satellite imagery is negligible [79], and anthropogenic structures are not masked by the presence of a dense vegetation canopy as in other biomes. Data acquisition and exploration site detection were therefore not impeded by these potential limitations.

The oil exploration site identification framework presented here is, however, not without limitation. Although the framework was successful in detecting the location of oil exploration sites in Algeria and Niger, it was unable to detect some of the reported oil extraction wells in the Niger region based on ground-truthed data (43% omission error; table 2 and figure 3). This may be attributed to both temporal and
likely not result in misdetections of temporary sites if the time period between consecutive images considered were sufficiently narrow. Individual oil flares are also not large (maximum 150 m in size), and thus may only cover a few Landsat image pixels. Smaller oil exploration sites may thus be difficult to distinguish from the surrounding desert, especially if not bordered by other associated landcover classes. This was indeed the case for the sites detected in the Niger study area, which (i) were smaller in size than those in the Algerian study area and (ii) could not additionally be identified by surrounding associated landcover classes, unlike the sites detected in the Algerian study area (i.e. water pools, settlement and roads). These results might also suggest that the oil flares class alone could be used in monitoring the distribution and expansion of oil exploration and extraction activities in desert regions over full landcover classification. However, we argue that due to the recent commencement and temporary nature of some of these activities in Niger, there may be no large surrounding structures yet constructed at these sites. As these become developed into the future, the applicability of classification of associated landcover types in locating and detecting their expansion will increase and will indeed be useful in monitoring expansion of these activities in other desert areas where sites are more fully developed (as is the case in Algeria). The SLC-off sensor issue with the Landsat 7 ETM+ data may result in missing data values over the locations of oil exploration sites in desert regions, resulting in the lack of detection of sites and suggesting a potential limitation of the use of Landsat 7 ETM+ data in these analyses. However, with the recent launch of the Landsat 8 satellite in early 2013 [80], data availability for the future implementation of this framework will not be hindered by this issue. Cross-scene atmospheric correction, which would serve to further normalize features across scenes [81], was not applied in these analyses and may have influenced the success of the implementation of this framework to the scene in Niger. However, owing to the applied methods using texture metrics to detect the limited infrastructure in the Niger scene and the few landcover classes present here, it is unlikely this correction would significantly affect the results of these analyses. Cross-scene correction as proposed by Furby & Campbell [82] should be considered in future implementation of this framework. Future and repeated implementation of the identification framework in other areas will assist in determining the severity of these potential limitations. Moreover, given the largely cloud-free satellite coverage of desert ecosystems [79], the application of multi-date change detection analyses across multiple scenes to assess and predict temporal changes in the distribution of oil extraction and exploration sites in these study sites may also help to further determine the relative success of our framework. The methodology employed herein is only one of the many options available to practitioners for the mapping and monitoring of oil exploration and oil exploitation activities in desert ecosystems. Other possible methods include the use of purely geographic information systems-based analyses, as employed by Osti et al. [34]. However, this form of analysis relies heavily on global energy and environment data providers and consultancies, such as the information handling services, which is not free of charge and may potentially be less objective than the use of open access satellite imagery. Alternatively, another potential remote sensing approach could be to use an active sensor, such as radar (e.g. TanDEM-X using synthetic aperture radar digital elevation model (12 × 12 m resolution)).

Figure 4. The thermal signature of pixels classified as the oil flares landcover class (through band 6–1 of the Landsat ETM+ sensor; [77]) converted to surface temperature shows temperature between 303 and 312 K across pixels of this class; panel (a) depicts the frequency of surface temperature (K) for pixels of the class oil flares (N = 2039) for the results of the framework classification in Algeria, and panel (b) depicts the values (N = 221) for the application area in Niger.

spatial limitations to our monitoring framework. Oil exploration activity, which had recently commenced in Niger at the time of these analyses [30,50], is often temporary in nature [33]. Indeed, four field-derived extraction sites in the Termit massif region (one within the study scene; figure 3) were known to no longer exist in November 2012. Accordingly, some of the non-detected sites in Niger, for which further field validation beyond 2011 does not exist, could have similarly been no longer in existence in the November 2012 scene. Thus, the relatively high omission rate for Niger may be due to non-detection of potentially temporary sites that were no longer present in the Landsat 7 ETM+ scene considered here. However, for continued monitoring purposes, the application of the framework to time-series Landsat satellite data would
Felbier et al. [83] demonstrated the application of this sensor in developing urban footprint maps (here human settlements), which identify anthropogenic structures with relative accuracy in desert regions. Interestingly, the spatial resolution of these data is very high, and the process is automated [83]. However, this method is targeted at settlements in general and would require additional information and methodological refinement in order to distinguish settlements from oil exploration sites, in addition to being costly and only implementable with the regularity of the data (current data availability covers only 2011 and 2012; [84]).

The relevance and requirement for a monitoring framework for oil exploration such as that developed here is further highlighted by the findings of this study. Ground-truthed data from SCF suggest that an overlap between the borders of the TTTNNR and oil extraction activity, and a conflict of interest between conservation efforts and oil exploration, already exists in Niger (figure 3). Moreover, while the identification framework employed here did not detect any exploration sites within the PA, many of the sites detected were less than 15 km from the border of the TTTNNR and the framework detected a novel site in close proximity to its boundary (figure 3), suggesting that potential and novel threats are encroaching into such areas. Increasing trends in oil prospection activities are similarly already a concern for the PA network across other global dryland regions, with concessions for oil exploration being sold within many PA boundaries [34]. Globally, the threats to dryland ecosystems from oil exploration and extraction activities are great [84–86], causing reductions in species abundance and biodiversity [87,88]. With much current conservation attention on the relative effectiveness of the global PA network for conservation efforts [89–95], methods to monitor threats to conservation and management action in PAs are of the utmost importance. If the expansion of oil exploration activities continues to be conducted at the edges of and within desert PAs, they have the potential to increase pressures on the animal populations within through increased edge effects [37,38], reducing the effectiveness of buffer zones around PAs and the resilience of animal populations and ecosystems within from impacts of further processes of land-use and climate change. Moreover, these threats have the potential to reduce the connectivity of the PA system to wider ecosystems [96], which may be of particular concern in desert ecosystems due to both the patchy distribution of essential resources related to the high variability of the rainsfalls in the Sahara [19,21], as well as the fragmented nature of desert biodiversity hotspots and highly endangered endemic species [14,18–21]. Indeed, increased anthropogenic activity within PAs and at their boundaries may create barriers to nomadic and migratory species’ movement [35,36]. The importance of maintaining desert biodiversity hotspots within PAs and their surroundings has never been more important, as global climate change is predicted to produce among the highest temperature increases in terrestrial biomes within these regions [23], highlighting severe potential impacts from anthropogenic water extraction associated with oil exploration activities. Water scarcity in and regions has been found to cause large reductions in species diversity and population size [21,96–98]. Resultant increases in water-related conflict between oil exploration activity and local biodiversity under future climate change in desert regions [23] could thus potentially lead to the collapse of already fragile populations (e.g. the red-fronted gazelle (Eudorcas rufifrons); [18]). Moreover, monitoring of the expansion of novel oil exploration across desert regions may be of great importance for determining sites for future reintroduction efforts for threatened desert animal species under climate change (e.g. the scimitar-horned oryx (Oryx dammah); [99]).

5. Conclusion

With mounting evidence of conflict between conservation efforts and energy procurement, a cost-effective monitoring framework for these activities is clearly of enormous relevance and utility in defining where adaptation and conservation actions must be targeted under future expansion, in order to secure the viability of unique biodiversity across global desert regions in the face of associated threats. This will become increasingly essential into the future, with projected increases for oil production and exports for many global dryland regions: Northern Africa (including many Saharo-Sahelian countries; i.e. Algeria, Egypt, Sudan), Central and Western Asia, Central America and Australia [100,101]. Under such future increases in hydrocarbon production, not only the utility but also the efficiency of the monitoring framework developed herein will become critical. Currently, the framework is not automated, and visual interpretation and data processing are required. Automation of the framework to include data acquisition, satellite image processing, and classification model generation and prediction across novel areas would serve to render the framework applicable across these widely spread geographical areas in an effective, non-labour intensive manner. Moreover, in the face of mounting conservation costs and the lack of funding targeted at global desert regions [15,16], the necessity for cheap and easily implementable monitoring frameworks for anthropogenic activities will become increasingly great into the future. We therefore highlight here the importance and relevance of the continued availability of open access satellite imagery, such as that from NASA’s recently launched Landsat 8 satellite, to on-the-ground conservation monitoring and management direction. Implementation and automation of the identification framework developed will enable the continued monitoring of areas of concern, advancing current understanding of extremely under-studied and -represented global desert regions, and forming a crucial resource in targeting on-the-ground management and conservation priorities in the face of global environmental change.

Acknowledgements. We would like to thank J. Reise and two anonymous reviewers for helpful comments on previous versions of this manuscript.

References


