



HAL
open science

PEPS Nouranat 2020 RICOCHET (A la Recherche de l'Intégration des Connaissances dans l'Observation des CHangements Environnementaux: mise en oeuvre d'une recherche-aTelier en Sibérie Orientale (Khamagatta))

Sébastien Gadat

► To cite this version:

Sébastien Gadat. PEPS Nouranat 2020 RICOCHET (A la Recherche de l'Intégration des Connaissances dans l'Observation des CHangements Environnementaux: mise en oeuvre d'une recherche-aTelier en Sibérie Orientale (Khamagatta)). [Rapport de recherche] CNRS UMR 7300 ESPACE; Aix-Marseille Université (AMU). 2021, pp.58. hal-03154446

HAL Id: hal-03154446

<https://hal.science/hal-03154446>

Submitted on 28 Feb 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License



**RAPPORT DE RECHERCHE
PEPS Nouranat 2020 RICOCHET
Février 2021**

Sébastien Gadal
sebastien.gadal@univ-amu.fr

Préambule

Objectifs

Le projet PEPS Nouranat 2020 RICOCHET (A la Recherche de l'Intégration des Connaissances dans l'Observation des CHangements Environnementaux : mise en œuvre d'une recherche-aTelier en Sibérie Orientale (Khamagatta)) avait pour objectif de constituer d'une part :

- (1) Une base de données géospatiale en Système d'information géographique (SIG) intégrant (1.1.) les enquêtes et cartographies mentales de perception des populations de Khamagatta aux risques systémiques et récurrents ; (1.2.) les analyse et modélisation par télédétection spatiale des processus en cours : érosion, feux de forêts, exposition aux risques d'inondations de débâcles et de glaces, extraction des structures des paysages (identification des paléo-chenaux, etc.).
- (2) Développer des modèles de simulation de l'exposition aux risques environnementaux des populations locales : crues de débâcles, feux de forêts, évolution des paysages de pergélisol.
- (3) D'analyser les processus d'adaptation des populations locales en vue d'une compréhension multidimensionnelle des dynamiques de changements environnementaux et sociaux au sein desquels les savoirs concrets et systèmes de mesures par télédétection sont parties intégrantes du système d'observation et d'analyse scientifiques.

Contexte

La pandémie de la Covid-19 et l'impossibilité de se rendre en Fédération de Russie et en Yakoutie a modifié l'organisation du programme de recherche. Les déplacements sur le terrain d'étude de Khamagatta de Madame Anne Tricot et Monsieur Jean-Louis Ballais ont été reportés en septembre et ces derniers ont ensuite fait part de leur souhait ne plus participer au programme de recherche, au regard, notamment, de la situation sanitaire. Au mois de septembre, contrairement à ceux portés par le coordinateur du projet et le doctorant impliqué, M. Zakharov, leur travail de recherche n'avait pas encore débuté. La composition de l'équipe de recherche a été de fait modifié.



Equipe de recherche

Les missions de terrain (validation des modèles développés par télédétection, analyse spatiale, et simulations), (enquêtes de terrain auprès des populations locales de K. et d'Amga), (les interviews auprès des maires des ulus), ont été menées par des collègues du département d'écologie et de géographie de l'université fédérale du Nord-Est -NEFU) sur les terrains durant plus de 11 semaines : *Moisei Zakharov*, doctorant AMU à l'UMR ESPACE (géographie physique et géomatique), *Antonina Savvinova*, Maître de conférences (SIG et géographie sociale), *Zhanna Degteva*, Maître de conférences (géographie humaine et sociale), *Yuri Danilov*, Professeur (géographie physique).

Piotr Janniec, doctorant NEFU en biogéographie et en télédétection a contribué à la création de modèles prédictifs d'exposition aux feux de forêts des populations locales ; *Jūratė Kamičaitytė*, Professeure à KTU, chercheuse associée à l'UMR ESPACE, sur l'analyse et la conceptualisation des perceptions des populations locales et l'aménagement.

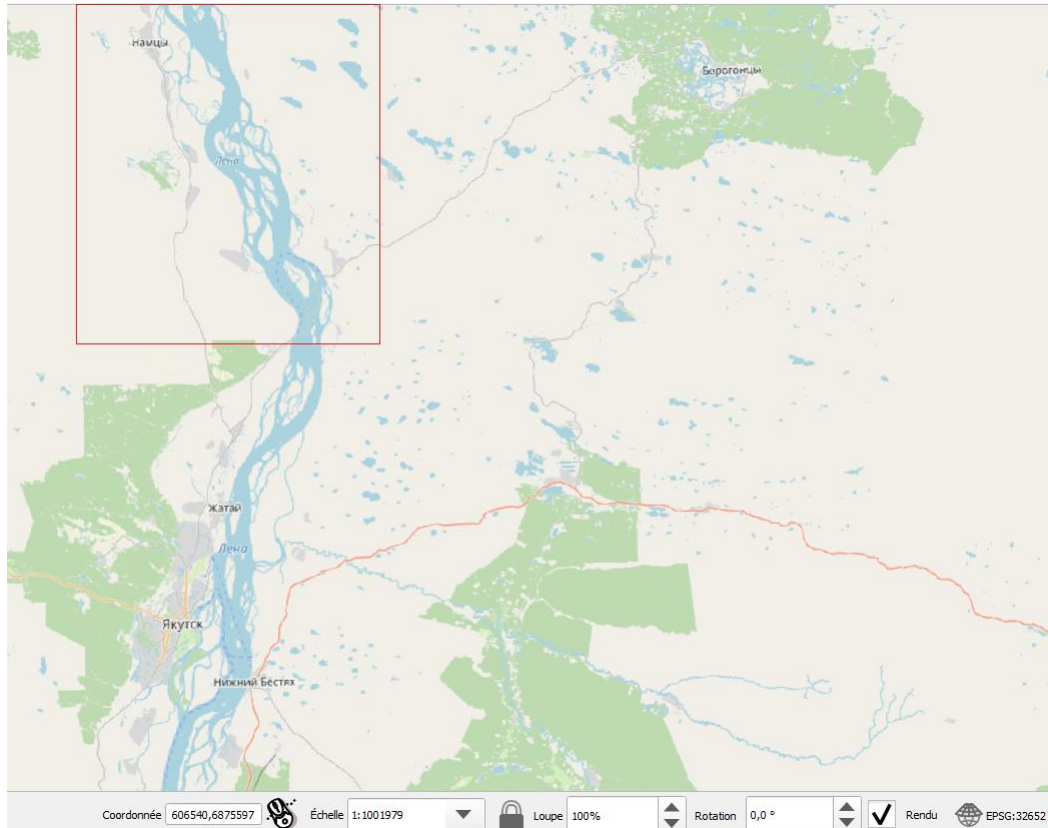
Le maire de Khamagatta, *Nikolay Zakarov*, et une enseignante de l'école franco-sakha de Khamagatta, *Sargylana Vasilieva*, ont été toujours présents pour nous aider et mener à bien et dans les meilleures conditions, malgré la pandémie de la Covid-19, les missions de validation terrain et les enquêtes auprès des habitants.

Les travaux et analyses des résultats, tout comme l'enrichissement de la base de données de perception des populations est en cours. Elle a vocation à rester active ses prochaines années, à évoluer, et à être déployée sur d'autres centres urbains. Elle se combine à une mise à jour annuelle ou biannuelle des spatio-cartes de l'occupation et de l'évolution des sols mesurant les impacts et changements qui en découlent. Également, à enrichir les modèles de simulation et d'exposition aux risques environnementaux des populations locales.

1. Introduction

Un des défis de la sécurité humaine dans les régions arctiques, et plus particulièrement dans les régions arctiques de la Sibérie orientale, concerne l'évaluation de l'exposition des habitants à des risques environnementaux récurrents pour la population urbaine dans un environnement géophysique caractérisé par des dynamiques de changement notables : dégradation du permafrost, d'inondations (inondations par les rivières, les canaux géomorphologiques et l'érosion du plateau) et de débâcle, d'érosion et de feux de forêt. Les centres urbains et les villages sont confrontés et exposés de facto à une multiplicité de risques naturels combinant dégradation du pergélisol, feux de forêts, inondations, etc. Cependant, l'évaluation de leur exposition est partielle ou inexistante ; d'où la nécessité de créer des observatoires mesurant l'exposition potentielle aux risques environnementaux des centres urbains. Ces derniers sont construits par l'analyse spatiale des données d'enquête sur la perception et les connaissances de la population locale et par la télédétection spatiale qui sont intégrées dans les systèmes d'information géographique. La connaissance et la compréhension de la façon dont la population locale perçoit l'environnement naturel et anthropique dans le contexte du changement climatique et de ses conséquences sont quelques-uns des facteurs clés pour créer des stratégies efficaces de planification urbaine pour les villes situées dans les zones à haut risque de risques environnementaux.

La région urbaine de Iakoutsk est la zone d'étude test pour la mise en place de ce type d'observatoire mesurant les évolutions combinant impacts de la variabilité climatique et urbanisation. Il contribue à la construction d'indicateurs sécurité humaine dans l'Arctique russe sibérien. Ce vaste territoire urbain unique localisé en zone subarctique et caractérisé par un climat arctique de plus de 370 000 habitants, située dans l'une des régions les plus froides du monde, est confrontée à l'accélération du réchauffement climatique et de l'urbanisation sur une base récurrente et saisonnière : l'urbanisation et le développement des centres urbains situés au Sud et au Nord de Iakoutsk (332 000 habitants) entre avril et septembre ; les crues de débâcles (eaux, glaces) co-suivi de l'érosion des berges de la Léna en mai, la dégradation du pergélisol sur les infrastructures, les lieux d'habitations, les zones industrielles et espaces agricoles, et les feux de forêt l'été. Le centre urbain de Khamagatta localisé à 70 km au nord de la métropole de Iakoutsk fait partie intégrante par ses interrelations journalière (pendulaires liées au travail). Il connaît depuis les années une croissance démographique continue passant de 1600 habitants en 2002 à 1800 habitants en 2020. Ce processus de croissance de la population est observé notamment par télédétection spatiale à partir des images satellites DMSP OLS-PL (capteurs visible proche infrarouge).



Khamagatta GIS, S. Gadal et M. Zakharov, 2020-2021

L'absence et l'obsolescence des données géographiques et cartographiques disponibles, exception faites des bases de données démographiques, nous ont poussé à combiner trois démarches d'analyses :

(1) la *télé-détection spatiale* avec le suivi de l'évolution de l'occupation et de l'utilisation des sols (Sébastien Gadal, Moisei Zakharov), des structures paysagères de pergélisol (Sébastien Gadal), des zones impactés par l'érosion des berges et du plateau (Moisei Zakharov), les feux de forêts ; à partir desquels sont développés des modèles de simulation spatiale des l'évolution des paysages, des modèles d'expositions aux risques de crues par exemple (Sébastien Gadal), et prédictifs sur les feux de forêts (Piotr Janniec, Sébastien Gadal).

(2) Les résultats sont validés ou non par les *enquêtes de validation terrain* (Moisei Zakharov), qui dans le cadre de ce programme de recherche, ont dû être menées par nos collègues de l'université fédérale du Nord-Est (NEFU) ;

(3) et enrichies par les *connaissances vernaculaires des habitants et populations locales*. Ces dernières sont transcrites sous formes de cartes mentales, d'analyse des résultats issus des traitements d'images satellites, et du questionnaire d'enquête mis en place (disponible en ligne : <https://forms.gle/hfgaTcfFaamEWQTj8>). Ce dernier a été mis en place par Antonina Savvinova, Moisei Zakharov, et Sébastien Gadal en combinant les retours d'expériences du programme ANR PUR (Polar Urban Centers) (WP-4) (Effremova et al., 2019) et du Peer-to-Peer Bilateral Dialogue Initiative 2015–2016 de l'US-State Government avec NEFU, the Northern (Arctic) Federal University (NARFU), et the University of Alaska Fairbanks (UAF) (Kontar et al., 2018).

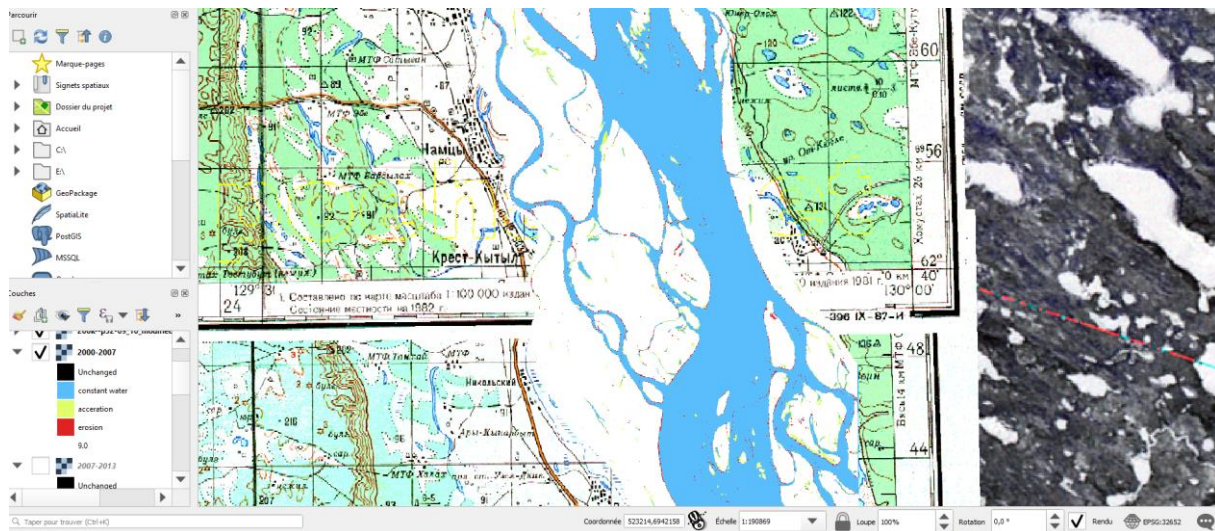
(4) L'intégration de l'ensemble des observations en un *système d'information géographique* (Moisei Zakharov, Sébastien Gadal).

Les Résultats acquis issus de l'observatoire de Khamagatta portent sur la création d'un SIG intégrant les bases de données créées à partir des images de télédétection spatiale et des enquêtes de terrain et en ligne auprès des habitants. Les résultats acquis issus de l'observatoire sont de x ordres :

(1) La création d'une base de données SIG sous QGIS intégrant des bases de données géoréférencées produites dans le cadre du programme de recherche :

- la base de données géoréférencée des cartes mentales et de perceptions intégrées en SIG.
- la base de données des enquêtes sur l'exposition aux risques naturels issu des enquêtes en ligne et de celles réalisées sur le terrain auprès des habitants. Elle est également intégrée en SIG.
- les spatio-cartographies de l'évolution de l'occupation des sols créées à partir des images de télédétection Sentinel 2A/B, Landsat 5 TM, et Landsat 8 OLI.
- les spatio-cartographies mesurant les processus d'érosions des berges des « bras » de la Léna.
- les spatio-cartographies des structures paysagères visant à extraire notamment les paléo-chenaux et hydro-géomorphologiques extraites à partir de Minimum noise fraction (MNF).
- les spatio-cartographies des Alas (lacs de pergélisols) combinant approche OBIA et base de données géolinguistique (issue du programme de recherche financé par la Russian Science Foundation « Land Ontology : semantics, semiotics, and geographic modeling », Sébastien Gadal, 2015-2018).
- les spatio-cartographies des feux de forêts.

Elle intègre également les bases de données démographiques, OSM, les bases de données Google Earth, et topographiques disponibles –



Khamagatta GIS, S. Gadal et M. Zakharov, 2020-2021

(2) Mise en place de modèles de reconnaissance des structures paysagères basées sur des modèles d'intelligence artificielle et ontologiques (Gadal et al. 2020) ; et prédictifs d'exposition des

populations aux feux de forêts basés sur des algorithmes d'apprentissage (machine learning) (Janniec, Gadal, 2020).

(3) L'intégration des connaissances vernaculaires des habitants de Khamagatta aux données de télédétection dans un premier modèle de connaissance géographique (Gadal et al. 2021-accepté). Modèle à partir duquel des indices de vulnérabilités peuvent-être générés (Gadal et al. 2021-soumis).

A partir des bases de connaissances et de données acquises structurées en SIG, les travaux de recherche portent actuellement sur :

(1) La création de modèle de connaissances socio-environnementales intégrés prenant en compte les mesures par télédétection spatiale, de terrain, et de perception. Elle se base notamment sur les travaux préalables menés à Iakoutsk et Kaunas en Lituanie (Gadal, Ourghemmi, 2019).

(2) L'analyse des résultats sur les modes et stratégies socio-territoriales d'adaptation et de résilience des habitants.

(4) Leur théorisation dans le cadre d'une démarche transdisciplinaire et l'aménagement (Kamiciuityte et al., 2020).

(3) Les stratégies d'aménagement ou de non-aménagement.

(4) La simulation de l'évolution des lits mineurs et majeurs à partir des résultats issus de l'extraction des structures paysagères par télédétection (avec notamment l'Ivannov Institute of System Programming RAS Moscou pour le développement d'un nouveau modèle stochastique).

(5) L'acquisition en continue des données de terrain, de perception des habitants, et de télédétection.

Les premiers résultats ont été publié ou soumis à publication dans :

1. Piotr Janiec, Sébastien Gadal. A Comparison of Two Machine Learning Classification Methods for Remote Sensing Predictive Modeling of the Forest Fire in the North-Eastern Siberia. *Remote Sensing*, MDPI, 2020, 12 (4157), pp.1-20. <10.3390/rs12244157>. <hal-03083192v2>
2. Sébastien Gadal, Moisei Zakharov, Jūratė Kamičaitytė, Yuri Danilov. Alas Landscape Modeling by Remote Sensing Image Analysis and Geographic Ontology. Study case of Central Yakutia (Russia). 6th International Conference on Geographic Information Systems Theory, Applications and Management, Polytechnic Institute of Setúbal/IPS; Knowledge Systems Institute; ATHENA Research & Innovation Information Technologies, May 2020, Online Streaming, Portugal. pp.112-118, <10.5220/0009569101120118>. <hal-02554659>
3. Sébastien Gadal, Moisei Zakharov, Jūratė Kamičaitytė, Antonina Savvinova, Yuri Danilov. Environmental Vulnerability Modeling in the Extensively Urbanized Arctic Center Integrating Remote Sensing, Landscape Mapping, and Local Knowledge. *vEGU 2021*, 19-30 April 2021, online, (accepté).
4. Sébastien Gadal, Moisei Zakharov, Jūratė Kamičaitytė. Human Perception and Environmental Risk Index Modelling with GIS in Yakutsk Urban Region (North-Eastern Siberia). *EUGE0 2021*, 26 June-1 July 2021, Prague (soumis).

2. Méthodologie

2.1. La nécessité de combiner données issues de la télédétection spatiale et connaissances locales

La connaissance locale de l'environnement social et physique est essentielle à la gestion des catastrophes naturelles (Tran et al., 2009). Les lacunes en matière de connaissances entravent sérieusement les efforts visant à renforcer la résilience aux catastrophes naturelles à tous les niveaux, en particulier dans les centres urbains de l'Arctique russe du Nord-Est exposés à de multiples expositions. Le manque d'information est le plus important au niveau local, en particulier en ce qui concerne l'information spatiale sur les risques, les ressources et l'organisation spatiale de la communauté (Liu et al., 2018). Les habitants sont bien conscients de la réalité géographique et peuvent indiquer quelles zones sont sujettes à des phénomènes jugés négatifs, y compris pour identifier les zones les plus inquiétantes à leurs « yeux ». C'est ce genre de connaissances locales qui permet aux scientifiques et aux décideurs d'explorer rapidement les besoins et les possibilités d'atténuation. Les points de vue transverse des résidents d'une collectivité urbaine comme Khamagatta - y compris les aînés et les élus en charge de l'ulus, les résidents et les enfants - rassemblent des points de vue et des préoccupations différents qui aident à cartographier les conditions jugées et pensées comme dangereuses au niveau local. A titre d'exemple, les lieux où l'expansion de l'urbanisation doit être limitée en raison des risques d'érosion ou d'inondations.

Nous avons, pour intégrer l'ensemble des connaissances issues de la télédétection spatiale et des connaissances locales, développé une approche méthodologique générale qui combine les processus cartographiques avec la participation des populations locales, la modélisation du paysage basée sur les données de télédétection.

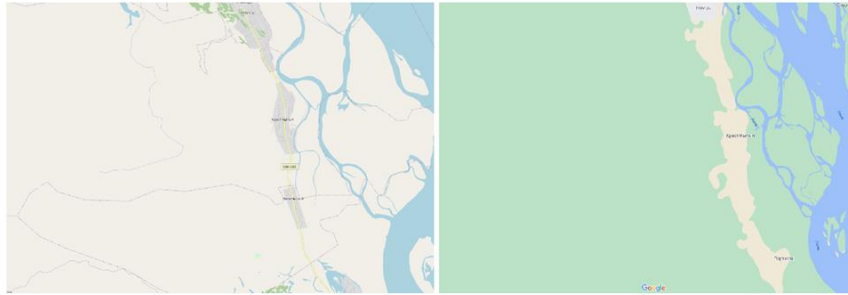
2.2. Cadre méthodologique général

La démarche choisie et développée se divise en deux approches parallèles qui sont ensuite intégrées en système d'information géographique :

-L'acquisition et la structuration des connaissances locales des habitants en SIG ; ce dernier étant enrichi par les données géographiques disponibles : Open Street Map, Google Earth, 2GIS, cartes topographiques scannées.

-Le développement de spatio-cartes, modèles d'évolution de l'occupation des sols, etc., à partir des images de télédétection spatiale Landsat 5 TM, Landsat 8 OLI, et Sentinel 2A/B.

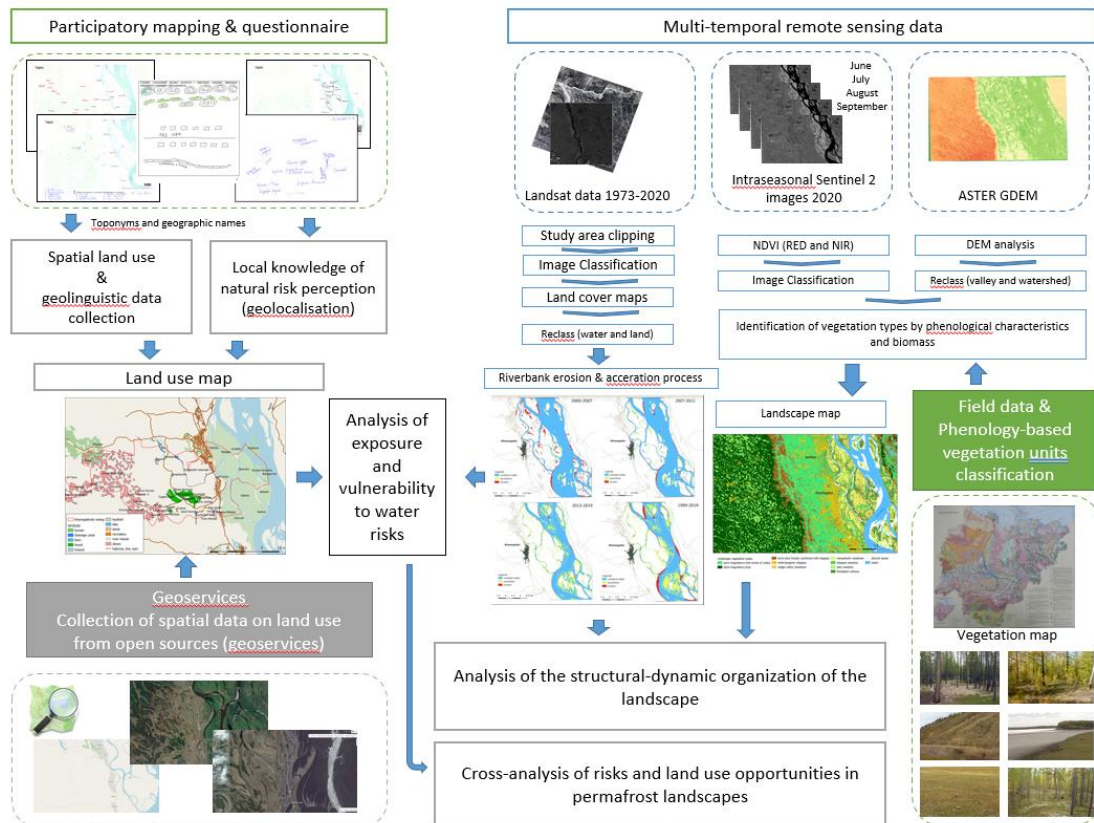
L'utilisation des images satellites est d'autant plus importante que les bases de données existantes ont des niveaux de précision et de qualités différentes, et ne prennent pas en compte les dynamiques de changements d'occupation des sols.



Open street map (a), 01.10.2020

Google map (b), 01.10.2020

D'un point de vue épistémologique, l'approche holistique choisie combine toutes les composantes de la complexité de l'espace géographique et répond à la fragmentation des connaissances (Lugeri et al., 2019). Cette démarche interdisciplinaire combine le traitement des données de télédétection pour la cartographie du paysage (Gadal et al., 2020), l'analyse spatio-temporelle de l'évolution des berges et des paysages, ainsi qu'une base de données géoréférencée intégrée en SIG des résultats de la cartographie participative, des questionnaires et de la collecte de données spatiales sur le terrain. La méthodologie générale s'organise telle que :



S. Gadal et M. Zakharov, 2020-2021

2.3. Acquisition et intégration des connaissances locales par approche participative

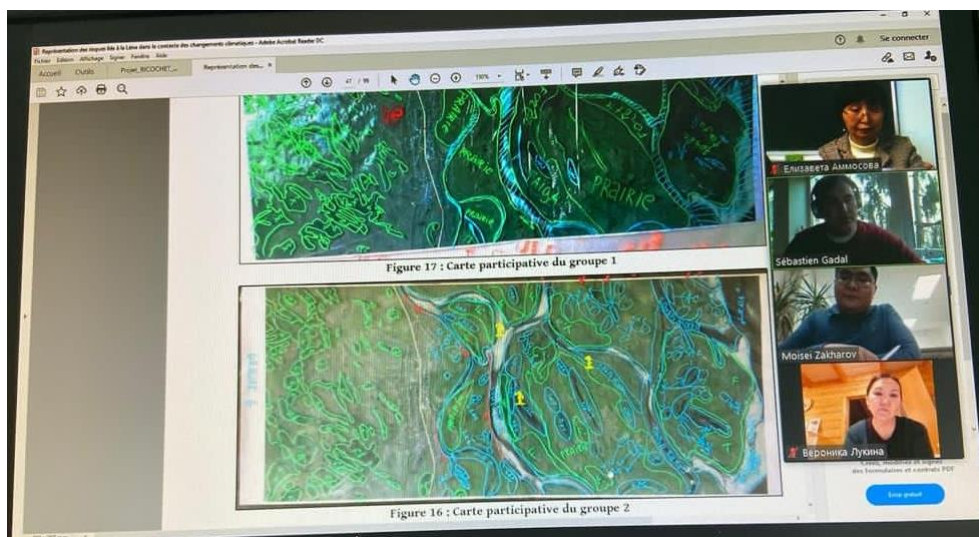
L'acquisition des connaissances locales s'est opérée sur 11 semaines de terrain via des enquêtes, un atelier de cartographie participative organisé en novembre 2020 sur place, des réunions virtuelles via Zoom chaque semaine entre octobre et décembre 2020 avec les habitants, et l'acquisition des données d'enquêtes via le questionnaire en ligne <https://forms.gle/hfgaTcfFaamEWQTj8>.

Ce dernier a été construit afin d'étudier la perception de la population locale de leur espace géographique vécu et perçu, d'éventuels phénomènes (environnementaux) (naturels) perçus comme négatifs. Par les événements naturels « négatifs » l'ensemble des répondants (98 personnes) ont répondu à l'unanimité (100% des réponses) : inondations printanières de printemps provenant de la débâcle de la Léna, pour 70%, les incendies de forêt et la fumée générée, l'érosion fluviale, les changements dans le littoral des îles de la rivière Léna pour 58 % d'entre eux.

L'ensemble de la population de Khamagatta ayant répondu au questionnaire en ligne a été victime d'inondations dans le village au cours des 20 dernières années en moyenne 15 à 16 fois.

Les résultats des enquêtes sont intégrés en SIG. Les trois ateliers de cartographies participatives qui ont été mis en place ont permis de recueillir et d'intégrer en SIG la structure et l'organisation de l'espace géographique vécu, de géolocaliser les zones touchées par les inondations, les incendies de forêts ou l'érosion.

Le premier atelier de cartographie participative a été organisé avec les enseignants et les élèves de l'école franco-sakha de Khamagatta en octobre 2020.

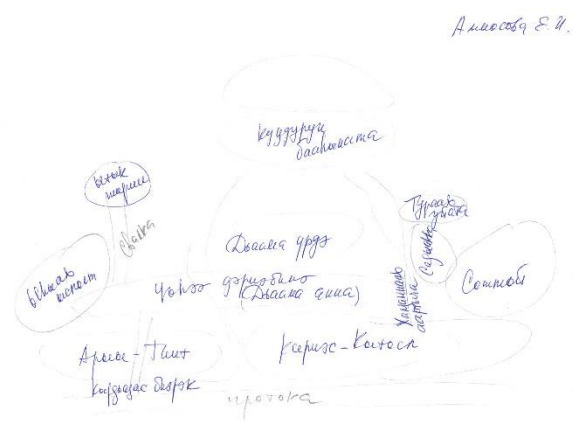
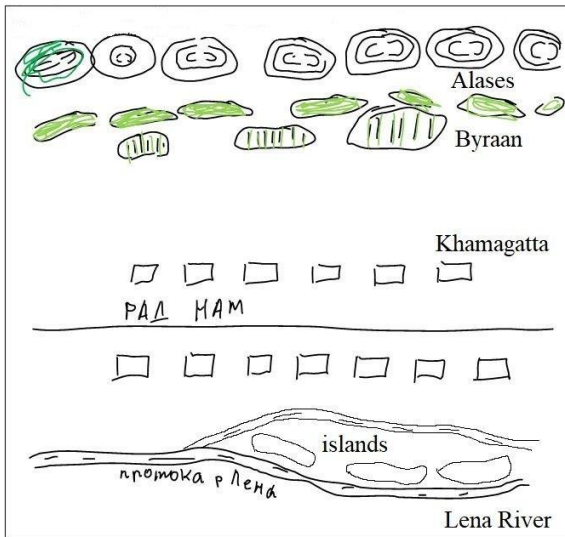


Analyse participative des résultats à partir des cartes créées par les enfants de l'école franco-sakha, octobre 2020



Réunion participative hebdomadaire avec les habitants de Khamagatta, novembre 2020

Le second atelier de cartographie participative organisé en novembre 2020 à Khamagatta a rassemblé 17 habitants (10 femmes et 7 hommes). Trois (sont) étaient employés de la municipalité, un (est) était un arpenteur foncier. Les participants à l'enquête sont des résidents de longue date du village, ont des parcelles subsidiaire (bétail, chevaux et champs de pommes de terre). Tous ont participé à des travaux de lutte contre les inondations chaque année : construction de digues anti-glace, évacuation du bétail, etc.). Plusieurs d'entre eux ont par des dessins ou l'écrit une conceptualisé l'organisation spatiale de leur territoire.



Exemples de représentation de l'organisation du Nasleg par les habitants, 2020.

Un troisième atelier de cartographie participative visait avec les mêmes habitants de Khamagatta à analyser les résultats des spatio-cartographies produites et à les enrichir via une application QGIS

(QField) sur smartphone, en utilisant les fonds de carte Yandex Maps téléchargées sur le site : <https://yandex.ru/maps> et des cartes topographiques scannées. L'ensemble des cartes participatives sont associées à un QR code unique. Tous les participants ont une connaissance précise et vécue de leur territoire avec la quasi-totalité des objets géographiques et de leurs emplacements qui organisent le Nasleg (Nasleg : c'est la plus petite forme d'une municipalité dans la division administrative-territoriale de la République de Sakha Yakutia). Nous avons demandé aux participants de tracer sur les cartes Yandex Map les objets géographiques liés à l'utilisation et à l'occupation des sols qu'ils jugent important. Ces derniers renvoient aux activités agricoles (par exemple les champs de foin (khoduya en langue sakha), aux structures d'habitats saisonniers (letniks : sailyks en langue sakha), mais également à la topographie (par exemple byraany : pente de la vallée de la Léna), à l'occupation des sols « naturels » (îles et zones forestières de taïga, etc.). Ceci a permis de corriger et de mettre à jour les cartes topographiques soviétiques à une échelle de 1 : 200 000 ; l'indexation des objets géographiques, leurs morphologies, et leurs toponymes. Des îles inexistantes sur les cartes topographiques et Yandex ont été indiqués (Ubas-Mummut a été corrigée sur Munur, l'île d'Eselyakh qui n'existait pas a été indiquée sur les nouvelles cartes). Il est fut de même pour les noms et toponymes des étendues forestières et de taïga. Certains objets géographiques ont une durée de vie annuelle (îles, etc.).

Les résultats des trois ateliers participatifs ont permis de mettre à jour les cartes et bases de données géographiques, et d'évaluer qualitativement les changements d'occupation des sols.



Exemple de cartographies participatives produites par les habitants, 2020.

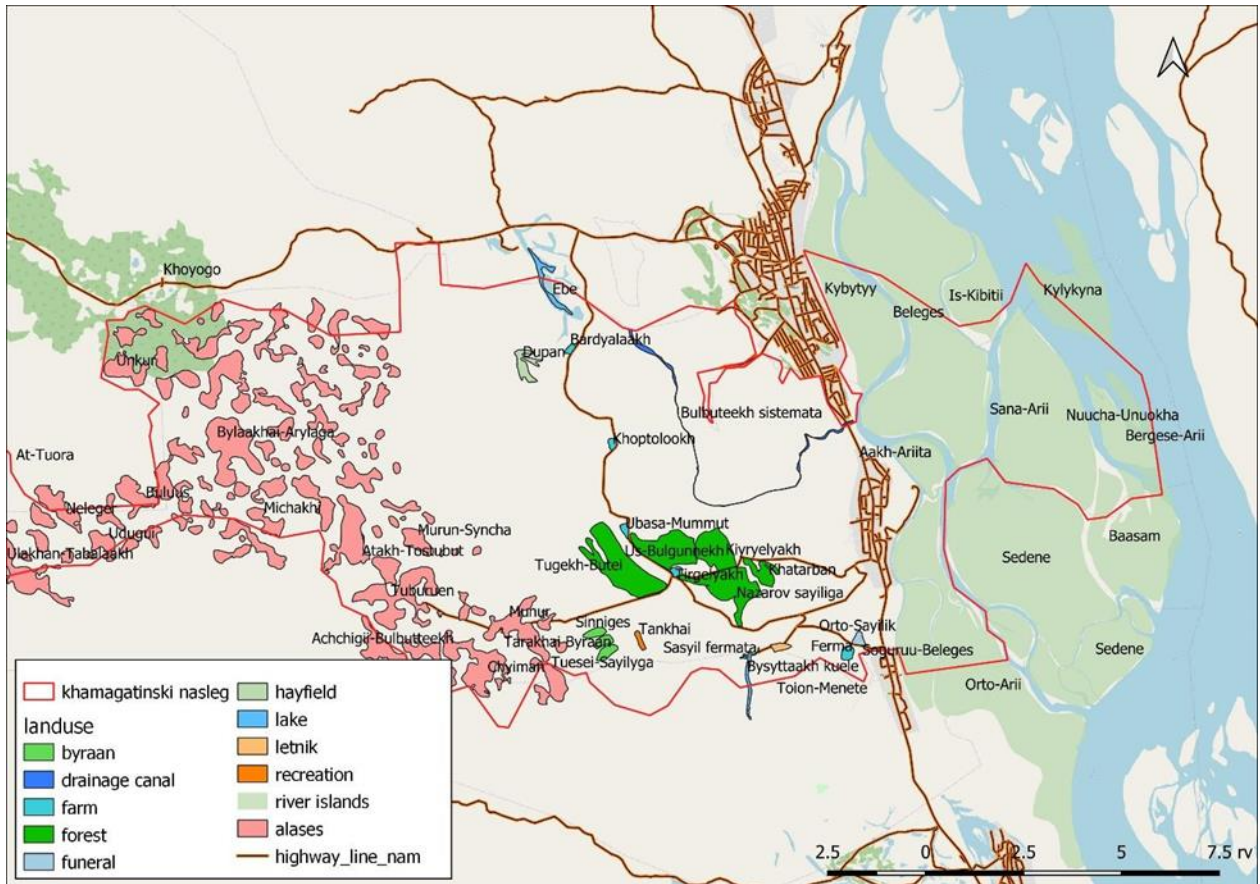


2.4. Création et analyse des cartographies participatives

Les résultats des ateliers de cartographie collaborative indiquent que les habitants de Khamagatta connaissent, localisent et orientent mieux îles voisines à proximité du centre urbain, que les paysages « naturels » environnements. Cela tient au fait que les îles de la Léna abritent presque toutes les terres agricoles des fermes locales (terres arables, pâturages, champs de foin et points d'eau pour le bétail). La totalité des zones boisées le long des routes ou dans les vallées sont mentionnées et cartographiées. Cela est dû à la présence de plantes sauvages recueillies par les résidents locaux.

En général, les participants connaissent bien la partie Sud du Nasleg. Moins concernant la partie Nord. Jusqu'en 2016, elle n'a pas été exploitée pour l'agriculture ou les activités économiques. Mais depuis l'application de la loi fédérale sur les « hectares d'Extrême-Orient », les habitants locaux ont acquis des parcelles de terrain dans ce territoire et leur culture suivra dans un proche avenir. La loi fédérale est en vigueur depuis le 1er juin 2016 et régleme les « relations » foncières, forestières et autres types d'occupations foncières mis à disposition des citoyens de la Fédération de Russie. Ces dernières sont généralement la propriété de l'Etat fédéral ou des municipalités sur le territoire district fédéral d'Extrême-Orient auquel est rattachée la République de Sakha.

Les informations et connaissances des participants sont ensuite intégrées en SIG et associées à la base de données Open Street Map (OSM), topographiques et aux données cadastrales pour générer la cartographie suivante :



M. Zakharov, S. Gadal, 2020, carte d'utilisation des sols synthétisant l'ensemble des cartes participatives.

Elle constitue une information géographique importante pour l'analyse des risques. Elle géolocalise les objets géographiques qui ont une valeur matérielle et/ou spirituelle ; ainsi que les activités économiques qui sont assujetties à des risques naturels.

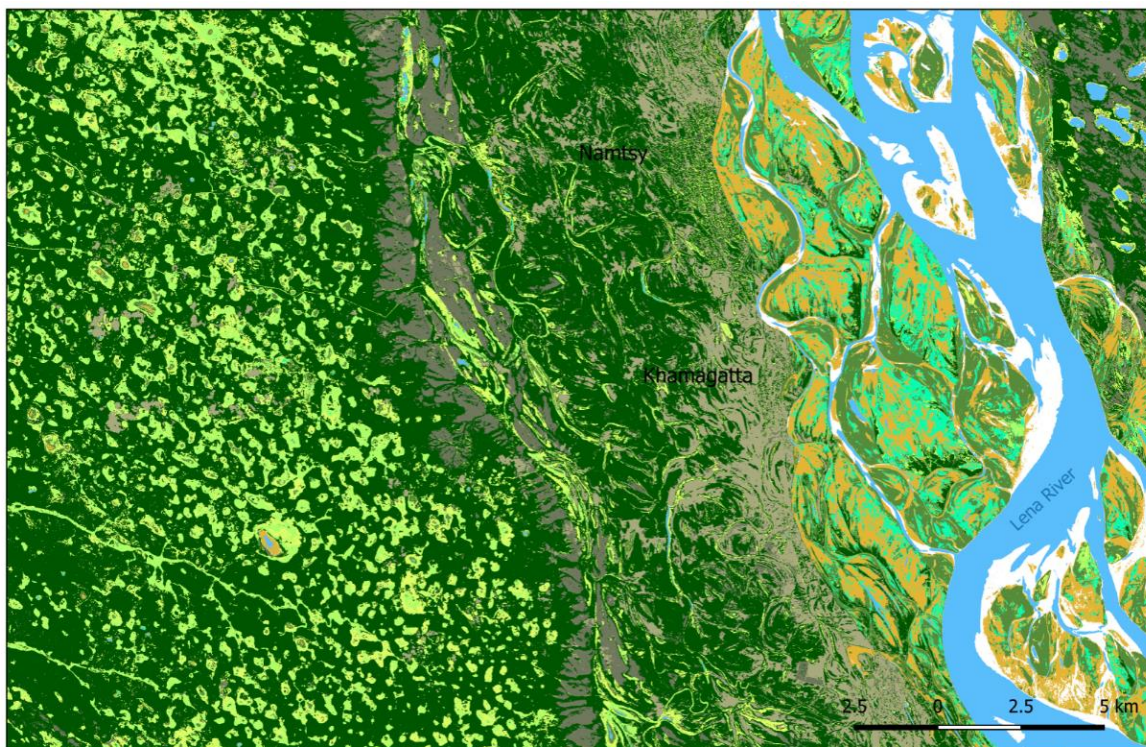
3. Approche par télédétection spatiale

3.1. Analyse inter saisonnière de l'occupation des sols à partir d'images Sentinel 2A/B

L'un des éléments les plus importants de l'évaluation de la vulnérabilité des territoires et de l'exposition aux risques naturels des populations est la cartographie du paysage basée sur la modélisation multicritères par les unités typologiques des paysages sur une échelle de pas moins de 1 : 200 000. À cette échelle, le principal indicateur de l'identification du paysage est les classes des associations de végétation, qui peuvent être interprétées à l'aide d'une approche de classification fondée sur la phénologie de l'analyse de l'image (Simonetti et al., 2014).

Les données d'entrée - pour la classification faisant appel à un algorithme d'apprentissage (*machine learning*) SAM (Spectral Angle Mapper) - est le NDVI (Normalized Difference Vegetation Index). Ce dernier est calculé à partir des bandes spectrales 4 et 8 de Sentinel 2 à une résolution spatiale de

10x10 mètres. Au regard de l'évolution phénologique du paysage, une classification des typologies du paysage a été effectuée en mai, juillet, et en août 2020 (absence de nuages). On notera que l'activité photosynthétique de la végétation s'étale de la mi-mai à la mi-septembre en laktoutie centrale. Pour chacune des dates, 8 classes de paysages de pergélisol sont reconnues, identifiées et cartographiées : (1) les paysages forestiers de mélèze (*Lárix cajánderi*) recouvert de lingonberries et d'herbassés ; (2a) les paysages forestiers de mélèzes et de pins (*Pínus sylvéstris*) avec des zones de végétation de steppe. (2b) On notera qu'une forêt a été identifiée sur les pentes bien drainées de la rive du substratum rocheux. (3) Les steppes anthropiques prévalent à proximité des zones d'habitations, là où la charge anthropique et le surpâturage sont accrus. Des prairies de steppe se développent sur le bassin versant. (4) les zones agricoles dans les vallées sont utilisés principalement uniquement comme letniks (saylyks en sakha), c'est-à-dire, pâturages de bétail en été. Elles sont caractérisées par l'éloignement géographique par rapport aux centres urbains et à la faible bio-productivité par rapport aux prairies de plaine inondable des îles de la Léna. (5) Sur les îles, il existe deux types de prairies qui sont importantes dans l'agriculture : (5a) les prairies inondables et (5b) les prairies méso-phytiques. (6) Les Alas (Gadal et al. 2020).



landscapes	Alas and sedge valley meadows	Willows floodplain
Larch lingonberry-forb forest	Steppe meadows	Alluvial sand
Larch-pine forests combined with steppes	Floodplain meadows	Water
Anthropogenic steppes	Mesophytic meadows	

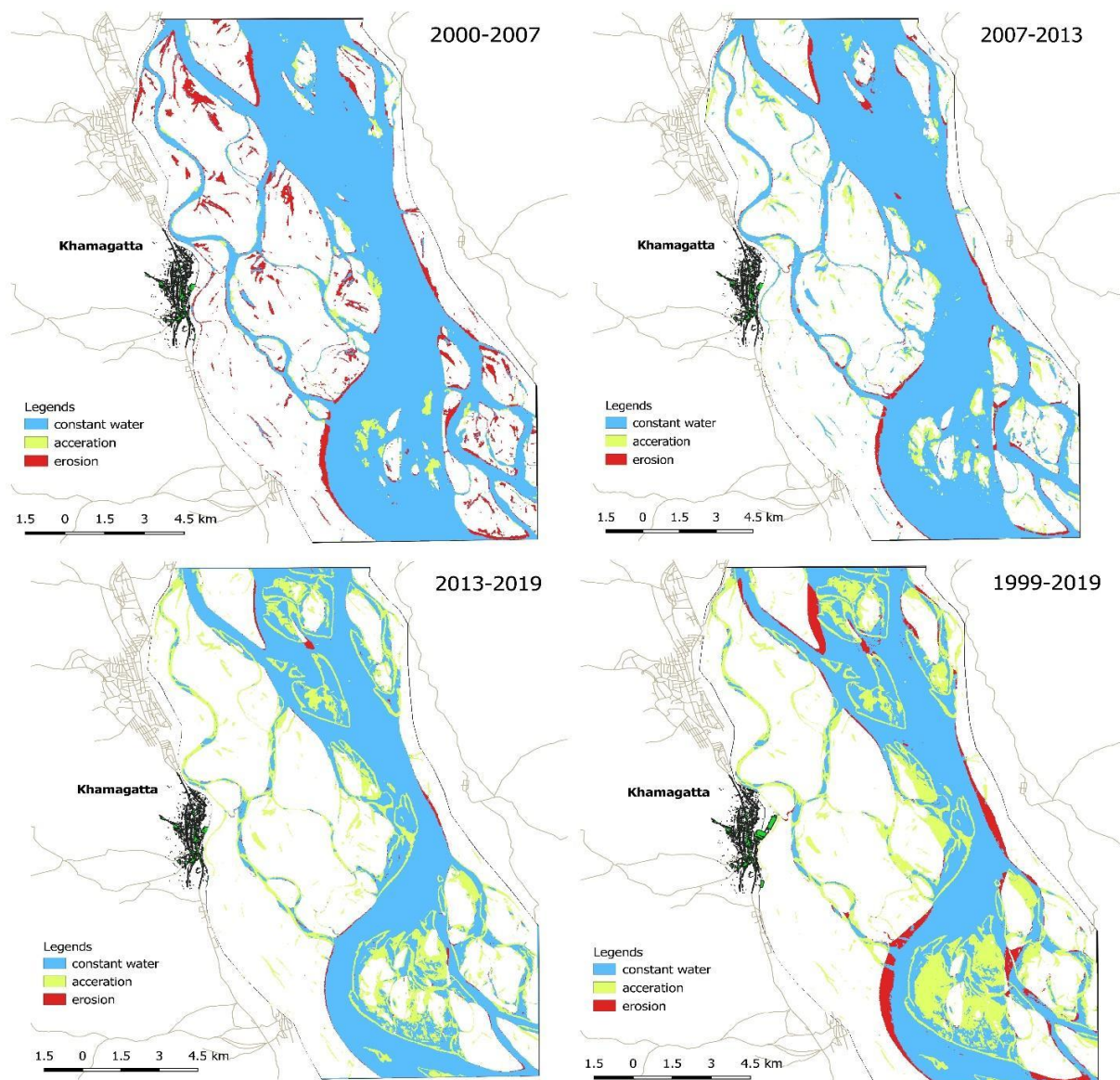
M. Zakharov et S. Gadal, 2020, spatiocartographie normalisée des paysages (mai, juillet, août 2020)

3.2. Suivi et modélisation de l'érosion des berges à partir des images Landsat 1 MSS, Landsat 5 TM, Landsat 7 ETM+, Landsat 8 OLI et Sentinel 2A/B

En référence aux résultats du questionnaire, l'érosion fluviale des berges est perçue négativement par 58% des habitants ayant répondu. Cela est principalement dû à la menace de destruction des terres littorales et agricoles situées sur les îles de la Léna. L'étude et la modélisation des processus d'érosion-accrétion sur les berges et sur les îles du centre urbain de Khamagatta a été réalisée à l'aide de données de télédétection, d'une carte topographique soviétique (P-52-IX). X) au 1 : 200 000 datant de 1980, de données de terrain et issues des ateliers de cartographie participative.

	Satellites	Bandes spectrales / résolutions spatiales	Dates
1	Sentinel 2 B	2,3,4,8 (10x10m)	Juillet 2020
2	Sentinel 2 A	2,3,4,8 (10x10m)	Juillet 2019
3	Sentinel 2 A	2,3,4,8 (10x10m)	Juin 2018
4	Sentinel 2 A	2,3,4,8 (10x10m)	Juin 2016
5	Landsat 8 OLI	2,3,4,5,6,7 (30x30m)	Juillet 2013
6	Landsat 5 TM	1,2,3,4 (30x30m)	Juillet 2009
7	Landsat 5 TM	1,2,3,4 (30x30m)	Juillet 2007
8	Landsat 7 ETM+	1,2,3,4 (30x30m)	Août 2000
9	Landsat 7 ETM+	1,2,3,4 (30x30m)	Août 1999
10	Landsat 5 TM	1,2,3,4 (30x30m)	Janvier 1989
11	Landsat 1 MSS	1,2,3 (120x120m)	Mars 1973

Le suivi et la modélisation des processus de modification des berges de la Léna ont été étudiés sur quatre périodes : 2000-2007, 2007-2013, 2013-2019 et 1999-2019. Elles ont été reconnues, identifiées et extraites par des algorithmes d'apprentissage automatique de maximum probabilité et Spectral Angle Mapper (SAM).



M. Zakharov et S. Gadal, 2020. Modélisation de l'évolution des berges de la Léna

Le volume total de l'érosion et d'accélération pour la période d'étude de 1999 à 2019 sur les deux rives du fleuve était de 112,2 ha et 188,7 ha, respectivement. Au cours de cette période, le taux total d'érosion et d'accélération est de 5,6 ha/an et de 9,4 ha/an, respectivement.

De 2000 à 2019, 58 hectares ont été érodés sur l'ensemble du territoire de Khamagatta (rive gauche à l'ouest). Néanmoins, 112 hectares ont été adjoints, avec une augmentation nette de 54 hectares de nouvelles terres.

De même, de 2007 à 2013 et de 2013 à 2019, l'accumulation a été supérieure à l'érosion. La quantité totale d'érosion au cours de cette période était d'environ 27,6 ha, comparativement à une augmentation d'environ 30,1 ha.

Entre 2000 et 2007, l'érosion a été plus prononcée que l'accrétion. Le volume total d'érosion et d'accrétion au cours de cette période était de 37,7 ha et 15,4 ha, respectivement.

Mais au cours des 20 années de la période d'étude, l'érosion totale des berges est inférieure au volume de l'augmentation. Les plaines inondables économiquement actives (plaines agricoles) et les prairies mésophytiques sont exposées à l'érosion des berges de la Léna, et la croissance est due aux fourrés de saule.

Le processus d'érosion cause plus de dommages et pourrait, dans un proche avenir, entraîner des pertes plus importantes de prairies précieuses utilisées pour la récolte du foin pour le bétail. Cependant, le résultat de l'érosion et de l'accrétion est influencé par certains facteurs naturels, tels que les dépôts de sable des rivières, le niveau de l'eau, le climat, etc. Ces facteurs créent des fluctuations de l'érosion-accrétion de la rivière.

Les changements de berges causés par l'érosion se produisent pendant le stade d'inondation en recul, lorsque l'excès de sédiments est déposé sous forme de bancs de sable dans les chenaux, ce qui entraîne un changement dans la direction du débit et de la migration des bancs de sable le long d'entre eux, ainsi qu'en raison d'une transgression du flux de matériaux.

Les résultats de cette étude ont montré que l'érosion, en particulier le long de la rive gauche de la côte, est la principale menace pour les îles de Sedene et d'Ortho-Ary ; tout comme pour les terres agricoles qui s'y trouvent. Elles corroborent les réponses au questionnaire ainsi que les cartographies participatives produites par les habitants de Khamagatta.

3.3. Modélisation et prédiction de la localisation des feux de forêts

Les résultats de l'enquête ont montré que l'exposition aux incendies de forêts étaient une préoccupation majeure pour 70% des participants. Elle est corroboré par les résultats de l'atelier effectué avec les élèves de l'école franco-sakha de Khamagatta.

La base de données FIRMS de la NASA récence (<https://firms.modaps.eosdis.nasa.gov/>) les feux de forêt à l'échelle régionale. Il nous a semblé intéressant au regard des travaux sur leur télédétection et caractérisation (Janiec et al., 2019) de développer à partir d'algorithmes d'apprentissage (machine learning) un modèle régional et un modèle local de prédiction et d'exposition des incendies de forêts en zone subarctiques et arctiques ; avec en premier lieu Khamagatta et la lakoutie centrale.

La copie de l'article ci-joint présente les deux modèles en annexe 1.

Piotr Janiec, Sébastien Gadal. A Comparison of Two Machine Learning Classification Methods for Remote Sensing Predictive Modeling of the Forest Fire in the North-Eastern Siberia. Remote Sensing, MDPI, 2020, 12 (4157), pp.1-20.10.3390/rs12244157. hal-03083192v2



Références bibliographiques

- Effremova V., Danilov Y., Gadal S. 2019. Затонные наводнения на реке лена на территории городского округа Г. Якутск. Meridian, Издательство ИП Осенина Ирина Леонтьевна, « Холод как преимущество. Города и криолитозона: традиции, инновации, креативность » ("Cold as an advantage. Cities and Cryolitozone: traditions, innovations, creativity"), 20 (2), pp.113-118.
- Gadal S., Zakharov M., Kamičaitytė J., Danilov Y. 2020. Alas Landscape Modeling by Remote Sensing Image Analysis and Geographic Ontology. Study case of Central Yakutia (Russia). 6th International Conference on Geographic Information Systems Theory, Applications and Management, Polytechnic Institute of Setúbal/IPS; Knowledge Systems Institute; ATHENA Research & Innovation Information Technologies, pp.112-118.
- Janiec P., Gadal S., Ivanova S., 2019, Геоинформационное моделирование риска лесных пожаров в Республике Саха (Якутия). Успехи современного естествознания (Advances in Current Natural Sciences), 2019, pp.37-42. (10.17513/use.37237).
- Janiec P., Gadal S. 2020. A Comparison of Two Machine Learning Classification Methods for Remote Sensing Predictive Modeling of the Forest Fire in the North-Eastern Siberia. Remote Sensing, MDPI, 12 (4157), pp.1-20. (10.3390/rs12244157).
- Kontar, Yekaterina Y.; Eichelberger, John C.; Gavrilyeva, Tuyara N.; Filippova, Viktoria V.; Savvinova, Antonina N.; Tananaev, Nikita I.; Trainor, Sarah F. 2018. "Springtime Flood Risk Reduction in Rural Arctic: A Comparative Study of Interior Alaska, United States and Central Yakutia, Russia" Geosciences 8, no. 3: 90. <https://doi.org/10.3390/geosciences8030090>
- Liu, W.; Dugar, S.; McCallum, I.; Thapa, G.; See, L.; Khadka, P.; Budhathoki, N.; Brown, S.; Mechler, R.; Fritz, S.; Shakya, P. 2018. Integrated Participatory and Collaborative Risk Mapping for Enhancing Disaster Resilience. ISPRS Int. J. Geo-Inf., 7, 68. <https://doi.org/10.3390/ijgi7020068>
- Lugeri F. R., Farabollini P., and Lugeri N. 2019. Landscape analysis as a tool for risk reduction AIMS Geosciences Volume 5, Issue 3, 617–630.
- Simonetti E., Simonetti D., Preatoni D. 2014 Phenology-based land cover classification using Landsat 8 time series. Institute for Environment and Sustainability. Luxembourg: Publications Office of the European Union.
- Tran P., Shaw R., Chantry G., Norton J. 2009. GIS and local knowledge in disaster management: a case study of flood risk mapping in Viet Nam Epub.

Annexe 1

Article

A Comparison of Two Machine Learning Classification Methods for Remote Sensing Predictive Modeling of the Forest Fire in the North-Eastern Siberia

Piotr Janiec ^{1,2} and Sébastien Gadal ^{1,2,*}

¹ Aix-Marseille University, CNRS, ESPACE UMR 7300, University Nice Sophia Antipolis, Avignon University, 13545 Aix-en-Provence, France; piotr.janiec@etu.univ-amu.fr

² Ecology Geographical Department, Institute of Natural Sciences, North-Eastern Federal University, Belinskogo 58, Yakutsk 677000, Russia

* Correspondence: sebastien.gadal@univ-amu.fr

Received: 27 November 2020; Accepted: 16 December 2020; Published: date

Abstract: The problem of forest fires in Yakutia is not as well studied as in other countries. Two methods of machine learning classifications were implemented to determine the risk of fire: MaxENT and random forest. The initial materials to define fire risk factors were satellite images and their products of various spatial and spectral resolution (Landsat TM, Modis TERRA, GMTED2010, VIIRS), vector data (OSM), and bioclimatic variables (WORLDCLIM). The results of the research showed a strong human influence on the risk in this region, despite the low population density. Anthropogenic factors showed a high correlation with the occurrence of wildfires, more than climatic or topographical factors. Other factors affect the risk of fires at the macroscale and microscale, which should be considered when modeling. The random forest method showed better results in the macroscale, however, the maximum entropy model was better in the microscale. The exclusion of variables that do not show a high correlation, does not always improve the modeling results. The random forest presence prediction model is a more accurate method and significantly reduces the risk territory. The reverse is the method of maximum entropy, which is not as accurate and classifies very large areas as endangered. Further study of this topic requires a clearer and conceptually developed approach to the application of remote sensing data. Therefore, this work makes sense to lay the foundations of the future, which is a completely automated fire risk assessment application in the Republic of Sakha. The results can be used in fire prophylactics and planning fire prevention. In the future, to determine the risk well, it is necessary to combine the obtained maps with the seasonal risk determined using indices (for example, the Nesterov index 1949) and the periodic dynamics of forest fires, which Isaev and Utkin studied in 1963. Such actions can help to build an application, with which it will be possible to determine the risk of wildfire and the spread of fire during extreme events.



Keywords: wildfires; MaxENT; random forest; risk modeling; GIS; multi-scale analysis; Yakutia; Arctic; Siberia

1. Introduction

The disturbance regime in the boreal forest is extremely variable. Every year in Siberia, millions of hectares of forest are burned. Forest fires are one of the main factors causing not only long-term, harmful changes in plant ecosystems, but also contribute to the deterioration of living conditions in society, especially in the event of wildfires. Taiga fire is a natural phenomenon. Fires determine the normal, ecological functioning of the forest in this region. Forest fires are an inseparable part of the natural cycle. After the fire, there are favorable conditions for the young generation of trees. Wildfires are important for the indigenous peoples of Siberia. The territory after the wildfire turns into a pasture and people harvest berries for winter stocks. Fires have deep consequences and issues to the transformation of the landscape in the specific context of permafrost. Fire in forest areas where permafrost is spread is considered as an important factor in modeling the surface and influencing geomorphological processes.

Boreal forest plays a big role in the circulation of carbon dioxide in the regional and global scale. Changes in fire regime and climate in this region have already started, and they have an impact on the carbon releasing dynamic not only during the fire, but also years after the fire when the permafrost is thawing [1,2]. Forest fires in Russia, in general, and in the Republic of Sakha (Yakutia) are one of the main and most common natural threats. Currently, the area covered by forest fires has increased significantly. According to the results of scientific research, not the least role in the spread of fires is played by the human factor. Almost half of all forest fires that have occurred in Yakutia are caused by humans [2]. The occurrence and spread of forest fires are influenced by a complex of different factors that mutually reinforce each other and create conditions that contribute to the forest fire.

Geoinformation systems and modern remote sensing methods are a popular and effective means of identifying the most important factors affecting fires. GIS technologies are effective for building the necessary long-term fire safety models in countries such as Canada, the USA, and most of the European countries, where sustainable forest management is wide. GIS forest fire risk models are commonly used, but most of them do not include the human factor [3]. In Russia, the Nesterov index, from the 1940s is still in use [4]. There is large interest in the development of an integrated fire hazard assessment system integrated with GIS, remote sensing, meteorological, and government data. A system like that, available for forestry, firefighters, and local authorities, can improve the fight against fires and reduce the damage, thanks to faster detection of fire ignition and focusing on the most endangered places. The purpose of this work was geoinformation modeling the long-term wildfire risk in the Republic of Sakha (Yakutia) on a macro- and microscale. To achieve this goal, the following tasks were defined: (1) study structure and regime of the wildfires in the boreal forest; (2) define the most important factors that cause forest fires in the region of the studies; (3) develop and implement statistical methods to determine the relationships between the wildfires and its factors; (4) develop spatial and statistical analysis methods for building a model of wildfire risk in macro- and microscale; and (5) validation of the built model and create wildfire risk maps at the macro-and microscale.

The methodology of the study of wildfire risk is based on forest management and geographical developments. The field of studies is built on Canadian, U.S., and European approaches to deal with

natural hazards as in the works of San-Miguel-Ayanz [3], Tedim [5], Oliveira [6], Cardille [7], Martinez [8], and others. Knowledge on the boreal forests in Yakutia was taken from the publications of such authors as A.P. Isaev [9], E.I. Troeva [2], M.M. Cherosov [10], and others.

Work was divided into two main stages. The first stage presents the methodology validating the impact of individual fire factors on fire occurrence. The result of the analysis is a determination of dependencies between factors and exclusion from the further classification of factors that do not affect fires. In the second step, based on the literature analysis [11–14] two methods of machine learning classifications were implemented to determine the risk of fire: MaxENT and random forest. The initial materials to define fire risk factors were satellite images and their products of various spatial and spectral resolution (Landsat TM, Modis TERRA, GMTED2010, VIIRS), vector data (OSM), and bioclimatic variables (WORLDCLIM).

The results in the form of long-term fire risk maps can be used for fire prevention and planning fire-fighting measures in the territory of the republic. The results can help create an application that can be used to determine the risk of fire and the spread of fire during a disaster. This work has the potential to lay the foundations for the future of a fully automated application of fire risk assessment in the Sakha Republic.

2. Material and Methods

2.1. Yakutia

To correctly understand fire regimes in Yakutia, it is important to know the features of their origin. The geographical, geological, climatological, and ecological position in the landscape make it possible to explain the complexity of the fire phenomenon. Yakutia is a very specific region, with occurrence extremely high and low temperatures, a thick layer of permafrost, specific geological structures, and the occurrence of light taiga forests dependent on fire regimes. Due to the high complexity of the territory, it was decided to study wildfires on two scales to examine the dependencies between fires and their factors on the regional and global scale. At the macroscale, the territory of research was the territory of the Sakha Republic; on a microscale, one of the regions of the republic was chosen, which was the Nyurbinskii district (Figure 1). This region was chosen due to the high number of fire incidents in recent years.

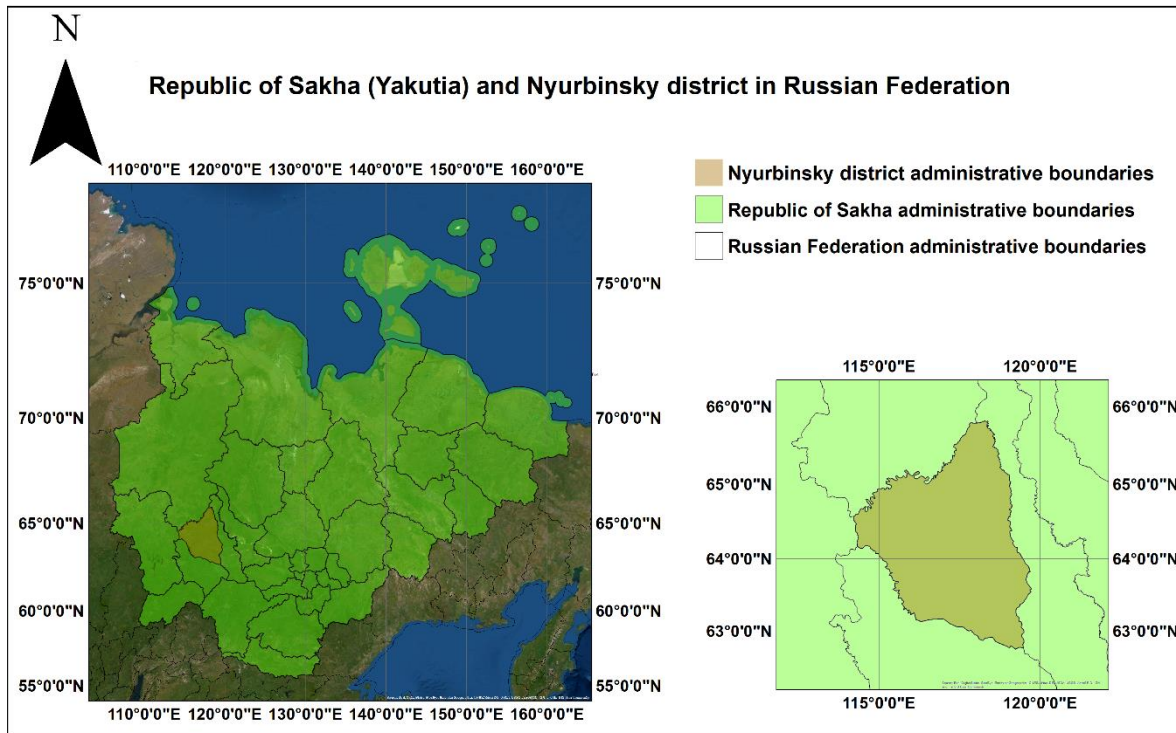


Figure 1. The geographical extent of the study area (source: Earthstar Geographics, low resolution 15 m imagery).

The Republic of Sakha (Yakutia) is in northeastern Siberia and occupies 1/5 of the territory of the Russian Federation. The territory of the republic from the east and south is closed by mountain ranges; in the north, it has access to the Arctic Ocean. The relief and geological structure are distinguished by a complex and diverse structure. The orography of the territory determines the characteristics of a sharply continental climate, permafrost, soil, vegetation, wildlife, grasslands, and grazing land, and influences the nature of human economic activity [15]. A characteristic feature of the climate of Yakutia is a sharp continentality, which is manifested in large annual fluctuations in temperature and a relatively small amount of precipitation. The main factor of this state is the Siberian anticyclone. In the study area, the average air temperature during the winter months is $-30\text{ }^{\circ}\text{C}$, and the average temperature of the coldest month (January) is $-35\text{ }^{\circ}\text{C}$. The average annual precipitation is 200–400 mm [16]. The permafrost thickness in the territory is sometimes more than 100 m. Seasonal thawing varies between 0.8 and 3.3 m, depending on the landscape and the type of soil. Permafrost in the conditions of Yakutia has an impact on all soil processes [17]. Sakha can be divided into three great vegetation belts. About 40% of Yakutia is located above the Arctic Circle, and all of this is covered with permafrost, which greatly affects the region’s ecology and limits the forests in the southern region. The arctic and subarctic tundra define a middle region where lichens and mosses grow like big green carpets and are favorite pastures for deer. In the southern part of the tundra belt along the rivers, scattered dwarf Siberian pines and larches grow. Below the tundra is a vast region of the taiga zone [2].

According to the state report “On the state and environmental protection of the Republic of Sakha (Yakutia) in 2018, the forested area occupies 82% of the territory of Yakutia, but the forest is 54% of the

territory. The area of the forest zone is 252,820.0 million hectares. The forest cover is very diverse, from 11.5% in the Verkhoyansk district to 91.7% in South Yakutia. In the Nyurbinsky district, the forest area is 4416.3 thousand hectares and is more than 84% of the total area. The timber reserved is about 9.2 milliard m³ and 96% of it is coniferous. The average timber reserve per one hectare is 58 m³. The biggest average timber reserves are of *Pinus Sibirica* stands (188 m³/ha) and *Picea spp.* (130 m³/ha). The average timber reserves of *Pinus sylvestris* is 104 m³/ha, *Larix spp.* is 62 m³/ha, and *Betula Pendula* 41 m³/ha.

A total of 95.6% of forest species in Yakutia are coniferous. The major species are *Larix* (4 main types: *Larix cajanderi*, *L. gmelinii*, *L. sibirica*, and *L. czekanowskii* (*L. sibirica* x *L. gmelinii*)), which represent 77.5% of the total forest resources. The second main species is *Pinus sylvestris* (6.5% of the total area) and *Pinus obovata* (0.24% of the total area). In southwest Yakutia, *Pinus sibirica* and *Abies sibirica* occur. *Picea ajanensis* is characteristic of the south mountainous regions. Other mountainous regions are occupied by *Pinus pumilo* (4.6% of the total forest area). The main deciduous tree species are *Betula pendula*, *B. pubescens*, *B. ermannii*, *Populus staveolens*, *P. tremula*, *Chosenia arbutifolia*, and *Salix spp.* Deciduous trees occupy only two million hectares, which are 1.24% of the total forest area.

Tree stands are well adapted to growing under extremely hard conditions of a dry climate and occurrence of permafrost. Forests are also very well adapted to the recovery process after fires. The process is determined by several factors. The most important factors are high seed production, good seed germination, and highly adaptive potential. *Larix cajanderi* is highly adapted to the existence with frequent fires. This species is pouring seeds in late summer in the ripening year. Seeds are being covered by needle litter and snow, where it creates the right conditions for sprouting and uses the spring moisture. The best conditions for forest growth are during the first years after a fire. Litter is destroyed, and the soil is enriched with ashy elements. The upper soil horizon's moisture is increased due to inflow from lower horizons [10].

Nyurbinsky ulus (region) is in the middle of the Vilyi River and occupies the territory adjacent to both Vilyi and its main tributary Marha. The region has an area of approximately 52.4 thousand sq km. Nyurbinsky region is part of the West Yakutia natural zone. It is characterized by plains and plateaus and the key elements are soil and vegetation on which relief has a big influence. According to Kurzhuev, the region is located on the Viluy Plateau, which is part of the Central Yakut Plateau. There is a characteristic occurrence of cryolithozone relief forms such as alases, yuryakhs, and bulgunyakhs. The number of days with snow cover is 210–225 days. The average annual temperature in the region is negative and in Markha it is –11.1 °C. The coldest month is January with an absolute minimum of –60 °C, and the warmest is July (the absolute maximum is +37 °C). The average amount of precipitation per year is 260–280 mm [15]. The maximum is in July. In the geobotanical regionalization of Yakutia, the Nyurbinsky region is situated in the boreal region in the taiga zone—a subzone of the middle taiga forests—sub-province, Central Yakutian middle taiga. The area of the region is dominated by *Larix* forests and *Pinus Sylvestris* forests. The river's valleys are characterized by rich meadows as well as common steppe and forest-steppe landscapes. Dry belts of alas vegetation are represented by the *Carex duriuscula* steppes [15].

2.2. Spatial Monitoring Factors for Fire Forest Monitoring in Yakutia

2.2.1. Fires Data

In the Republic of Sakha (Yakutia), open data on forest fires are not available. There is a database of forest fires at the Ministry of Environmental Protection (<https://minpriroda.sakha.gov.ru>), but these

data are tabular, and without georeferencing, it is impossible to create a GIS database necessary for this type of research. These data are understated due to the addition to the database of only those fires in which there was action taken to extinguish the fire. Fires that are far from human activity often do not extinguish, so are not classified in the database. The most popular global data source on fire data is the Fire Information for Resource Management System (FIRMS). Data available in the service are collected from the VIIRS 375 m, 750 m, and MODIS Collection 6 Active Fire Product. The data are collected from 2002 until now. For the studies, we chose data from the MODIS Collection 6 sensor because of their longer availability and enough spatial resolution. We used data between 2001 and 2018 from the FIRMS fire archive. The shapefiles were in the Geographic WGS84 projection. The confidence values ranged from 0% to 100% and ranged in one of three fire classes (low-confidence fire, nominal-confidence fire, or high-confidence fire) [18].

2.2.2. Factors Data

Forest fire regimes are extremely diverse and vary due to the spread of fires and climate changes [8]. In difficult to manage forest areas like in Yakutia, it is necessary to properly characterize the factors that are causing forest fires. Researchers have used several different variables to assess wildfire risk [3,19–22]. Due to the lack of data, it was decided to investigate only long-term fire factors data like constant, long-term factors such as slope, aspect, fuel, climate, NDVI, etc. Constant factors are those factors that do not change rapidly, but gradually, in the long perspective. Constant factors can be calculated in medium- or long-term periods before the fire season [3]. In Yakutia, there is a very big data shortage. If some data are already collected, it is very difficult to access them. For these reasons, mostly global datasets were used.

Variables were divided into four groups: (1) meteorological (precipitations, temperature, maximum summer temperature, radiation); (2) NDVI; (3) landform (elevation, slope, slope direction); and (4) human activity (distance from roads, distance from settlements, distance from rivers). All data types were in other formats and had different spatial resolutions (Table 1).

Table 1. Technical description of the data.

Factor	Source	Format	Resolution	Preprocessing
Radiation	WorldClim	GeoTiff	30 seconds	ROI
Precipitations				
Temperature				
Max temperature				
NDVI	MODIS/LANDSAT	Hdt/GeoTiff	250 m / 30m	Mosaic, resampling, atmospheric correction ROI
Elevation	GMTED2010	GeoTiff	7,5 seconds	Mosaic, resampling ROI
Slope				
Slope direction				
Distance from settlements	OSM	Shape		Eulicdean distance/ ROI
Distance from roads				
Distance from water lines				

To collect bioclimatic variables, the WorldClim dataset of global climate layers was used. The WorldClim dataset has a spatial resolution of about 1 km² [23]. WorldClim is a set of global climate layers (gridded climate data), specifically developed for ecological modeling on GIS. Currently, WorldClim provides several datasets for different temporal scenarios (past, current, and future conditions). In this work, data for the current condition scenario (1970–2000) were used. WorldClim bioclimatic variables are analysis-ready data so the preprocessing was not necessary.

NDVI is greatly used in the evaluation of the phenology and productivity of the vegetation. Onigemo et al. claims that the values of NDVI obtained through images at the peak of drought were related to the content moisture and fresh phytomass, showing its potential to estimate fire risk [46]. Illaera et al. found a good correlation between the NDVI values and the location of wildfires [24]. To assess NDVI in macro-and microscale, it was decided to use NDVI from two sources: MODIS and LANDSAT images. The Landsat data were used for the determination of NDVI in the Nyurbinsky region. As the region of interest, the following paths and rows were defined: 129-15; 129-16; 130-14; 130-15; 130-16; 131-15. There were selected available images from the year with the smallest cloud cover. Images without cloud cover at the peak of the growing season (July) were available only for Landsat 5 in 2009. Before creating the mosaic, atmospheric correction was necessary. It allowed for improving images from level 1 to level 2. For NDVI calculations, bands 3 and 4 were used by the formula described by J.W. Rouse in 1973 [25]. To determine the average NDVI in the vegetation season for Yakutia between 2001 and 2018 with MODIS, dataset MOD13A2 Version 6 was used. MOD13A2 provides NDVI values with a resolution of 1 km. The product is derived with a monthly interval. There were selected images at the peak of the growing season (July) from each year. At the territory of Yakutia, there are the following tiles, which were defined as region of interests: h21v01, h22v01, h23v01, h21v02, h22v02, h23v02, h24v02, h23v03, h24v03, h25v03. After the selection of the data yearly, an NDVI mosaic was created and the mean NDVI calculated between 2001 and 2015.

To model landform factor data, it was decided to use one of the most used in these types of works, digital elevation models: the Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010). It incorporates the current best available global elevation data. GMTED2010 is commonly used for radiometric and geometric correction, cover mapping, and extraction of drainage features for hydrologic modeling [26]. There was a necessity for mosaicking. For the GMTED, the following entities were used: GMTED2010N50E120, GMTED2010N50E150, GMTED2010N70E120, and GMTED2010N70E150. From the DEM, the information about elevation, slope, and slope direction was derived.

To investigate the relationship between the fire and the presence of a human, it was necessary to obtain information on the distance of fires from roads, buildings, and rivers. The data in shape format were downloaded from Open Street Map (OSM). According to previous works on this subject [53,54], to achieve this goal, the Euclidean distance was used. Euclidean distance is a straight line between two points in Euclidean space. The locations were converted to raster. The resolution of the raster was defined by the shortest of the width or height of the extent of the input feature, in the input spatial reference, divided by 250.

After the preprocessing, two regions of interest (ROI) were defined for each obtained raster. Each of them was clipped using the forest cover of the Nyurbinsky region and Republic of Sakha (Yakutia) forest cover. The forest cover was made available by the Institute of Biological Problems of Cryolithozone in Yakutia. The same ROI was made for fire points between 2001–2015 from the FIRMS dataset.

2.3. Modeling the Risk Assessment of Forest Fires

2.3.1. Fire Factors Analysis

The first step after the preprocessing (Figure 2) was the classification of the values in each raster, for ordering data and eliminating individual pixels strongly deviating from the average pixels were probably deviating due to measurement error or instrument inaccuracy. Due to a large amount of data, the surface to cover (3,084,000 km²), and computing capabilities, it was necessary to simplify the data to some extent, but retain their characteristics. Jenks's optimization method was used. The Jenks classification is designed to identify data clusters as well as maintain a representation of all data in the set. The optimization lasts as long as the limits of the intervals are obtained, respecting the principle of the smallest possible differentiation of the observations contained in them, with the greatest distance between the intervals at the same time. This method aims to reduce the variance within classes and maximize the variance between classes. This is done by striving to minimize the average deviations of each class from the middle class while maximizing the deviations of each class from the means of other groups [27]. New values in rasters have been extracted to each fire point from the FIRMS dataset of over 17 years in the territory of Yakutia and the Nyurbinsky region.

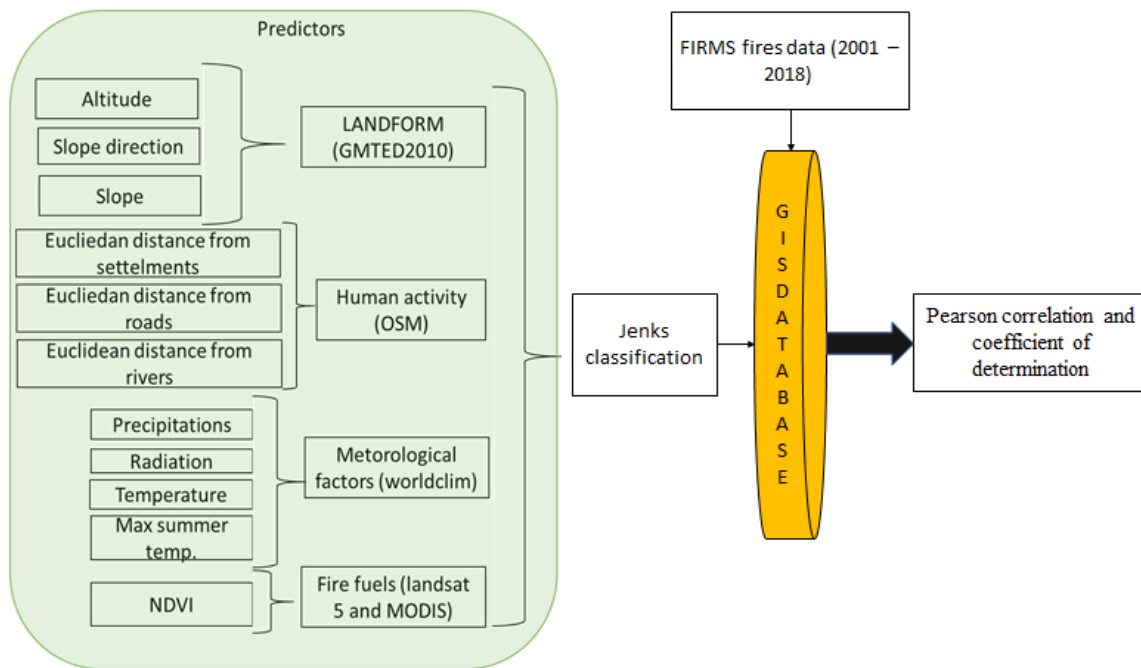


Figure 2. Schema of the analysis of the relationship between the fires and their factors.

The next step was to calculate the area of each class in each factor. Then, all fires between 2001–2018 were summed, separately for each category in each obtained earlier categorized raster. With these data, it was possible to determine the regression, and correlation between the dependent variables and explanatory variables. The same calculation scheme was repeated for each variable. In the last step, Pearson correlation and coefficient of determination between the fire data and factors data were calculated for each model.

2.3.2. Long-Term Fire Risk Modeling

The next step was to build risk models using two machine learning methods (Figure 3). The training dataset was built from the FIRMS fire data between 2001–2015. All registered fire points at the territory in the Republic of Sakha Yakutia and Nyurbinsky region in this period were used. The variables were divided, according to the previously obtained correlation coefficient. As the predictors, two types of datasets were examined. A dataset with a coefficient of correlation equal or higher than satisfactory (higher than 0.6) and a dataset with all 11 previously selected fire factors. All predictor rasters have been resampled and converted to grid format. Such a set of training data and predictor data were created, separately for the territory of Yakutia and the territory of the Nyurbinsky region.

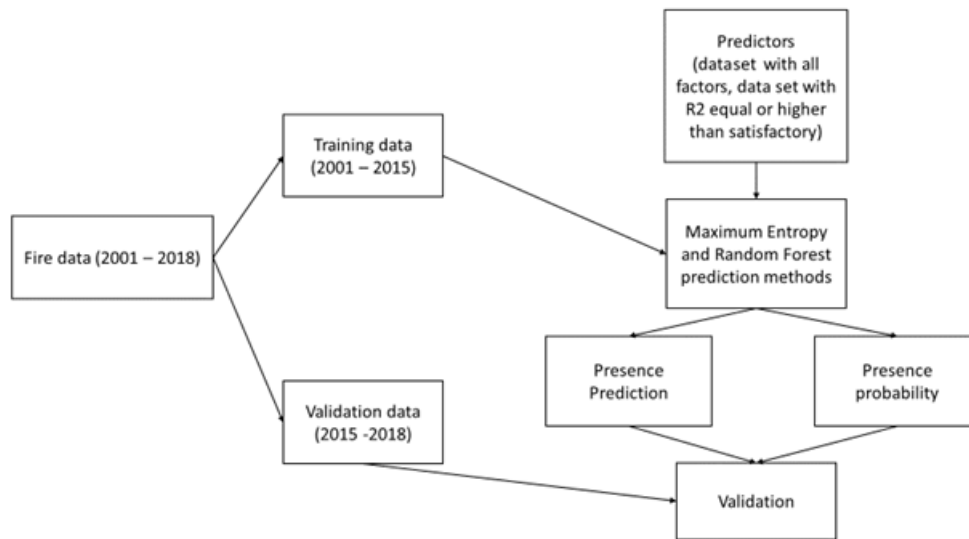


Figure 3. Long-term fire risk modeling methodology.

In the simulation, two types of prediction methods were selected, referring to the works of authors such as Peters [11], Parisien [28] (maximum entropy prediction model) and Oliveira [6] (random forest prediction model). In their studies, the authors showed a good correlation between the risk models obtained when using machine learning and forest fires. The authors demonstrated the superiority of methods using machine learning over traditional ones.

The maximum entropy prediction model is a widely used and accepted statistical method to produce predicting probability distributions. The model is adapted in diverse topics such as thermodynamics, economics, forensics, imaging technologies, and recently, ecology. Maximum entropy can provide accurate predictions of patterns in macroecology and help identify the mechanisms that matter most [29]. The algorithm is widely used for mapping species distributions [30] and conservation planning [31]. In recent years, the model of maximum entropy began to be used in the ecology of fires. Maximum entropy is a density estimation method based on a probability distribution. It is a presence-only machine learning algorithm that iteratively contrasts environmental predictor values at occurrence locations with those of a large background sample taken across the study area [13]. Maximum entropy has proved to be an enormously powerful tool for reconstructing images from many types of data [14].

The biggest advantage of a random forest is that it is a very flexible method, and it can be used in different types of problems. The random forest algorithm is a fully nonparametric machine learning method for data analysis. Classification and regression random forest is competitive with the best available methods and superior to most methods in common use [33]. The application of random forest can be an effective methodology to predict fire occurrence in different sites. The random forest algorithm is a technique developed by Breiman (2001) [33]. It combines a large set of decision trees, which is the biggest advantage compared to a simple decision tree algorithm. Each tree is trained by a set of variables, which are randomly selected from the training dataset.

All classifications were carried out using SAGA—6.4.0 and its modules “Maximum Entropy Presence Prediction”, “Random Forest Presence Prediction (ViGrA) Classification”. After the processing, two types of models were obtained: presence prediction and presence probability maps. Presence prediction maps are only meant to determine if there is a possibility of a fire. All raster pixels are classified into two classes:

Absence—there is the possibility of a forest fire.

Presence—there is no possibility of the forest fire.

Presence probability maps are meant to determine how high the possibility is of the fire. All raster pixels are classified into six classes:

Very low—very low possibility of the fire.

Low—low possibility of fire.

Moderate—moderate possibility of fire.

High—high possibility of fire.

Very high—very high possibility of fire.

Extreme—extreme possibility of fire.

To obtain only six possibility classes, it was necessary to reclassify the models. For reclassification, the natural breaks method was used.

For the validation of the results, raw fire points data between 2015 and 2018 in the Republic of Sakha (Yakutia) and the Nyurbinsky region were used. To carry out statistical analysis pixel count in each probability class in the presence probability maps and each class at the presence prediction maps were summed. Fire points in each class were summed and divided by the pixel sum in each class. This process was carried out to consider the surface of each class when verifying models. Next, the percentage share of fires in each class was calculated. The last step was regression and correlation analysis.

3. Results

3.1. Fire Factors

In the first part of the studies, attempts were made to uncover the main factors affecting the possibility of a wildfire. Of the 11 preselected factors, not all of them showed a good correlation with fire points (Table 2).

Table 2. Coefficient of correlation for each variable in the Republic of Sakha (Yakutia) and the Nyurbinsky region.

Factor	Yakutia	Nyurbinsky
Radiation	0,97	0,39
Precipitations	-0,33	-0,99
Temperature	-0,41	0,20
Max temperature	0,93	0,47
NDVI	0,91	0,97
DEM	-0,76	0,28
Slope	-0,97	-0,73
Slope direction	-0,43	0,15
Settlements	-0,97	0,86
Roads	-0,80	-0,82
Water lines	0,91	-0,54

In Yakutia (macroscale), the correlation above 70% was shown by factors such as solar radiation, maximum summer temperature, NDVI, elevation, slope, distance from roads, distance from settlements, and distance from rivers.

The problem was studied in the scale of the region using the example of the Nyurbinsky. In this example, only five factors out of 11 showed a good correlation. A very strong relationship between fires and precipitation was demonstrated. As the amount of precipitation falls, the number of points of ignition increases, which was not observed on a macroscale. The distribution of NDVI in microscale is almost the same as for the Yakutia. With an increase of the slope, fewer fires are observed, as in the macroscale.

3.2. Fire Risk

The next step was to create fire hazard models using two types of modeling methods. Fire risk models in the territory of Yakutia are shown in Figures 4 and 5. Results of the modeling differed, but both methods gave satisfying results. For the territory of the Republic of Sakha (Yakutia), high coefficients of correlation for each prediction method can be observed (Tables 3 and 4). Coefficients were slightly higher for the random forest prediction method. In this method, the model with 11 variables did not give considerably better results than the model with nine variables.

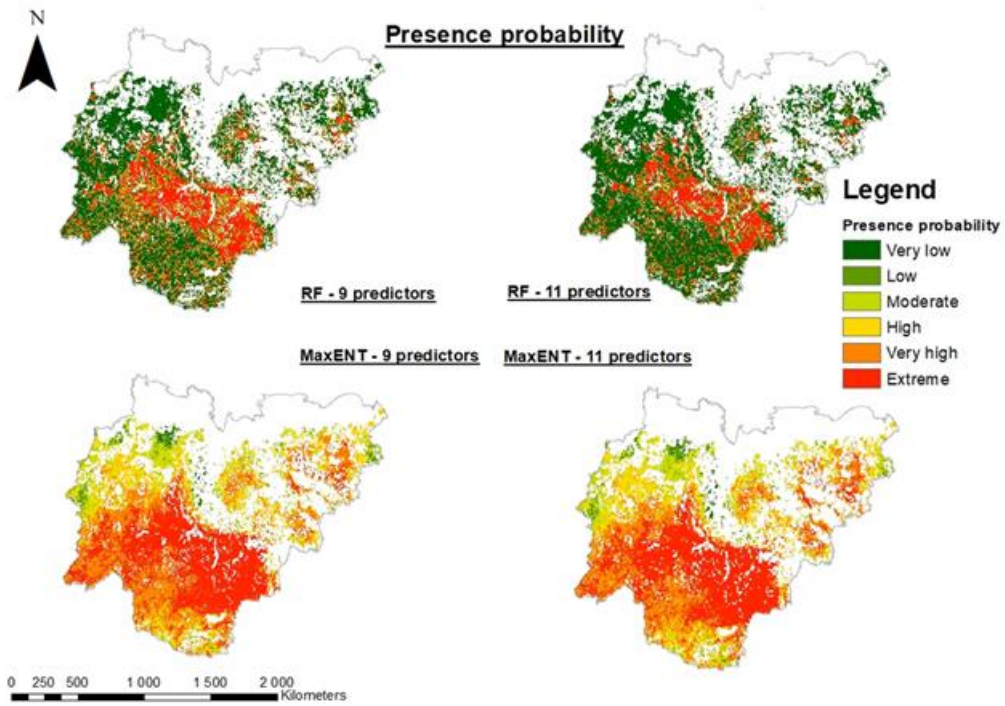


Figure 4. Results of the long-term wildfire presence probability in Yakutia (macroscale).

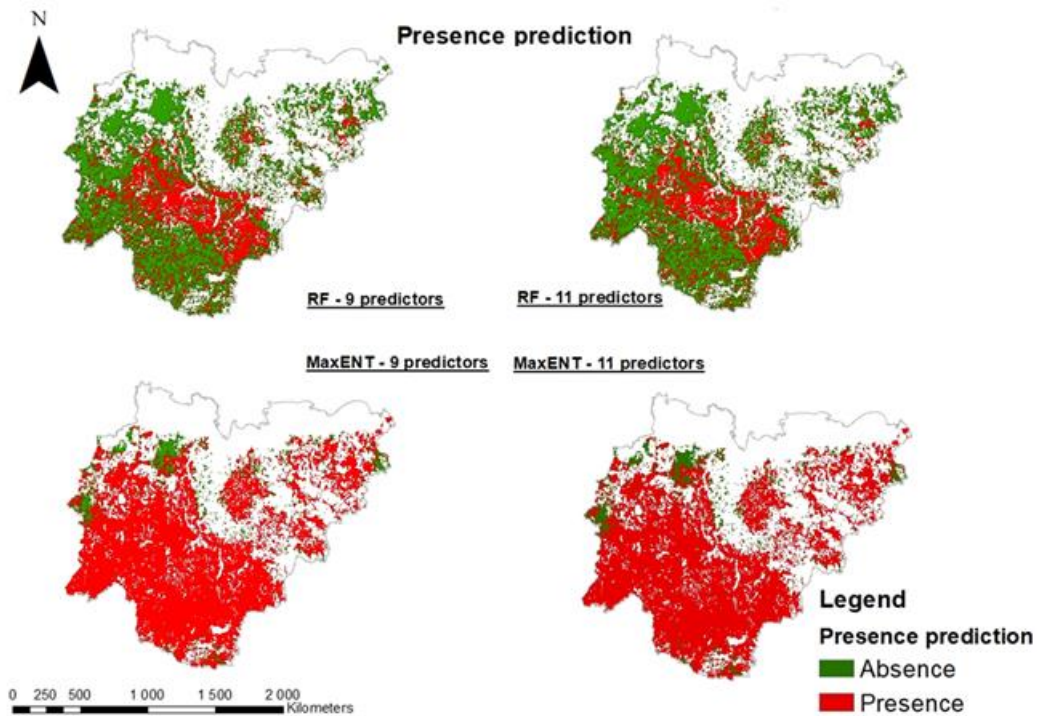


Figure 5. Results of long-term wildfire presence prediction modeling in Yakutia (macroscale).

Table 3. Comparison of the different types of long-term wildfire presence probability modeling methods in Yakutia (macroscale).

Fire probability	Percentage of fire points in each probability class			
	Random forest		MaxENT	
	9 Predictors	11 Predictors	9 Predictors	11 Predictors
Very Low	0,07	0,07	0,69	0,95
Low	1,74	1,65	1,57	2,30
Moderate	8,96	9,89	9,19	9,59
High	19,82	18,53	10,64	10,58
Very High	25,52	28,02	27,01	25,37
Extreme	43,89	41,84	50,89	51,21
R	0,97	0,98	0,91	0,90

Table 4. Comparison of the different types of long-term wildfire presence prediction modeling methods in Yakutia (macroscale).

Presence Prediction	Percentage of fire points in presence and absence class			
	Random forest		MaxENT	
	9 Predictors	11 Predictors	9 Predictors	11 Predictors
Absence	0,20	0,14	12,50	14,13
Presence	99,80	99,86	87,50	85,87

When modeling using the maximum entropy prediction model, better results gave a dataset with nine variables. In both models, the highest number of fires between 2015–2018 is in the high, very high, and extreme probability classes. In the very low-risk class, there were less than 1% of fire points. Presence prediction models showed a superior random forest method over the maximum entropy method. In the random forest method, almost 100% of the fire points were in the presence class. The maximum entropy method was characterized by worse results. In the absence, the class located more than 10% of the fire points from the validation dataset.

Modeling carried out at the territory of the Nyurbinsky region showed slightly different results than in Yakutia (Figures 6 and 7). Correlation coefficients were not as high and did not exceed 0.9 (Table 5). According to the coefficient of correlation, the smallest error was characterized by a model of maximum entropy using six predictor variables. The coefficient of correlation for 11 predictors was only slightly lower and was equal to 0.89.

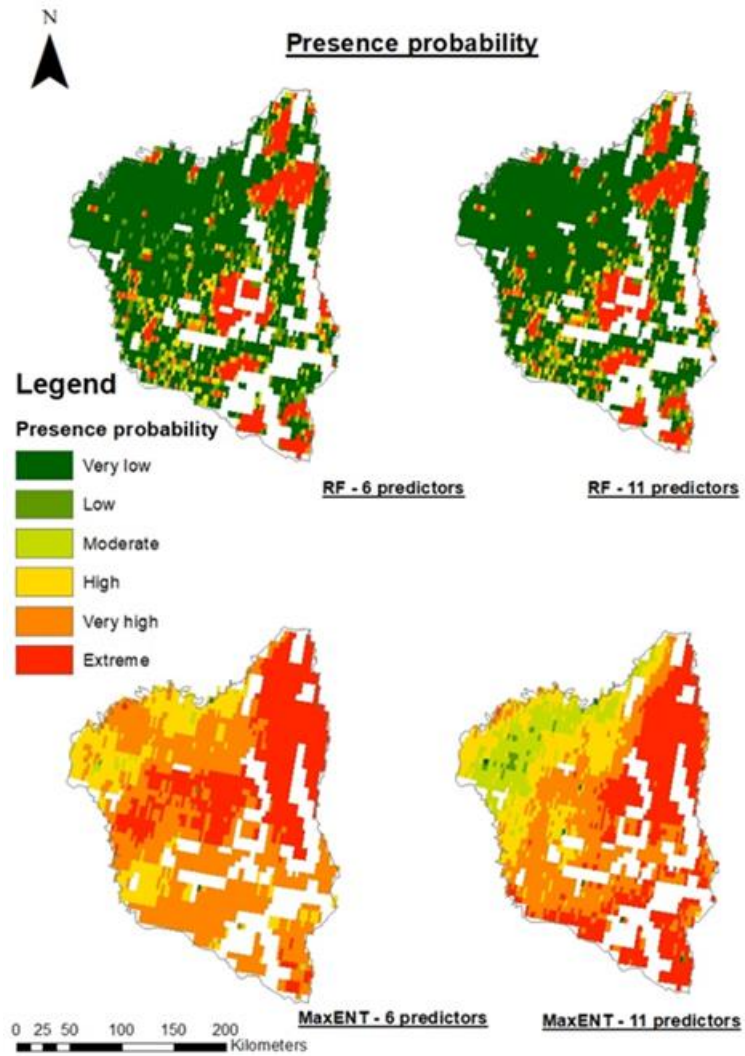


Figure 6. Results of the long-term wildfire presence probability modeling in Nyurbinsky (microscale).

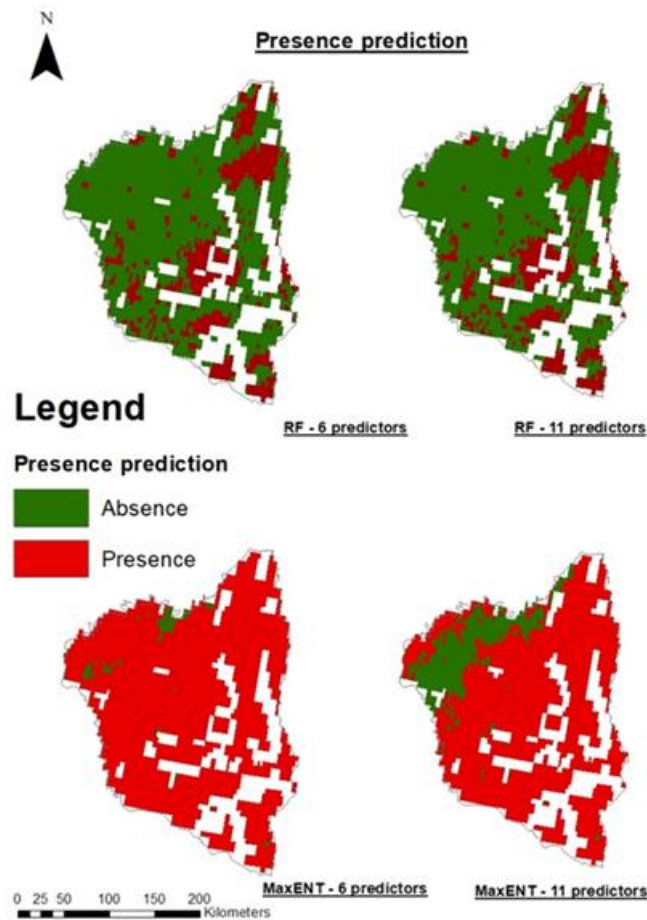


Figure 7. Results of the long-term wildfire presence prediction modeling in Nyurbinsky (macroscale).

Table 5. Comparison of the different types of long-term wildfire presence probability modeling methods in Nyurbinsky (microscale).

Fire probability	Percentage of fire points in each probability class			
	Random forest		MaxENT	
	6 Predictors	11 Predictors	6 Predictors	11 Predictors
Very Low	4,98	4,97	0,00	0,00
Low	15,09	17,04	0,00	0,00
Moderate	20,40	16,06	1,92	12,81
High	15,60	20,25	32,28	31,23
Very High	24,94	21,56	31,73	29,02
Extreme	19,00	20,13	34,07	26,94
R	0,75	0,82	0,90	0,89

It was observed that in the low and very low fire risk classes, there were no fire points. Significantly worse results were shown by the random forest prediction method. In both cases of the

random forest method (six and 11 predictors), the coefficient of correlation did not exceed 0.82. Over 20% of all fire points were at low-risk classes. Similar results were observed by analyzing the presence prediction maps (Table 6). The random forest prediction method was characterized by more than 30% of fire points in the absence class. The results were distributed in a similar way in the maximum entropy prediction model, which is characterized by 11 predictors. The best results were observed in the maximum entropy classification using six predictive variables. In this model, more than 90% of the points were in the presence class.

Table 6. Comparison of the different types of long-term wildfire presence prediction modeling methods in Nyurbinsky (microscale).

Presence Prediction	Percentage of fire points in presence and absence class			
	Random forest		MaxENT	
	6 Predictors	11 Predictors	6 Predictors	11 Predictors
Absence	32,73	31,22	8,34	32,66
Presence	67,27	68,78	91,66	67,34

It was observed that in both regions (Yakutia and Nyurbinsky), the random forest presence prediction method gave more accurate results. Even though the Nyurbinsky region incorrectly classified a larger number of fire points, it can be unambiguously stated that it is wrong. This is because (Figures 5 and 7) in the presence maximum entropy, presence prediction method, almost all of the territory of Yakutia and Nyurbinsky was classified as territory with the presence of fires. The results of the random forest presence prediction method showed a much narrower territory of the possibility of the occurrence of fires and yet it classified more points correctly in the territory of Yakutia.

4. Discussion

In fire studies, fire risk is one of the major topics and there are many different approaches to this subject. Blanchi et al., in their work about a methodological approach in fire risk studies, collected more than 50 works connected with fire probability cartography. Risk mapping has less than twenty years. Previously, there were rather preferred descriptive approaches. Since the 1990s, there has been an increasing interest in this field [41].

The study presented the possibilities of using different types of GIS and remote sensing data in modeling the wildfire risk. Results allow us to reflect various aspects of fire studies. The difficulties in fire risk assessment were to point out and clarify possibilities to define wildfire risk. There are many different approaches in hazard mapping based on different datasets, scales, and algorithms. The multi-approach is relevant in fire management studies [35,36]. Gai et al. [37] developed a spatially weighted index model for fire risk assessment. You et al. [38] integrated a Forest Resource Inventory Database based on four aspects of topographical, human activity, climate, and forest characteristic factors. Goldarag et al. [39] used neural networks and logistic regression for fire risk assessment. The biggest challenge in the study was to collect the necessary data due to poor exploration of the area of studies. GIS technologies in Yakutia are under development [40]. For this purpose, we chose a semi-probabilistic mode of modeling, which gave possibilities of combining historical fire data and physical fire mechanisms [41].

The results of the analysis showed that the fire risk assessment in the Republic of Sakha (Yakutia) is not a problem that can be easily solved. The specificity of the studied territory is significantly

different from other parts of the world that are facing the problem of wildfires. Boreal forests of Yakutia are perfectly adapted to extremely severe climatic conditions. Fires have been affecting the natural functioning of the forests of the region for centuries. Fires not only affect the climatic conditions, but also the formation of the terrain [42].

Results of the correlation analysis in macroscale showed that with increasing radiation, the risk of fires increases, and the same happened with increasing maximum summer temperature. Lim et al., in their study, highlighted that fire data showed a high correlation with climate factors [50]. NDVI was highly correlated with fires ($R = 0.91$). With increasing NDVI, the number of points of ignition increased. This situation was associated with the accumulation of combustible materials along with an increase in forest biomass. Elevation and slope correlated inversely proportional. According to Pourghasemi et al., slope and aspect are some of the most important factors controlling forest fire occurrence [48]. Kurbatsky et al. wrote that long drought conditions in the Yakutian valleys caused intense droughts and accumulation of fire fuels, which caused strong fires [34]. As the slope increased, we observed a decreasing number of wildfires. This situation may be due to the strongly inclined slopes that act as a barrier to the spread of fire. Additionally, on inclined slopes, many combustible materials cannot accumulate. Ghorbanzadeh et al. claimed that environmental variables are not the only reason for fire susceptibility and risk [52]. A very strong connection between the increase in the number of fire points and the human factors was observed. In places with proximity to transport routes, the number of fires increases, and the same thing happens if we consider the settlements. This situation indicates a strong anthropogenic impact on the risk of fire. There are not so many fires near water lines, which can be caused by the presence of moist areas close to the rivers.

The largest correlation at the microscale was observed when taking rainfall into account. As the rainfall increased, the number of fires increased. The other climatic factors did not show a strong correlation. NDVI, similar to the macroscale, strongly influences the number of fires. The properties of combustible materials relate to type, phytomass, condition, and moisture, among which the moisture content is the most important for fire protection [45]. In a completely different way, anthropogenic factors were distributed on the microscale than on the macroscale. The Nyurbinsky region has a relatively high population density compared to the entire territory of Yakutia, which may affect the differences between the results. Most fires continue to occur along roads, but the closer the settlements are, it occurs less frequently. This situation may be because fires that approach the villages pose a threat to people and fire-fighting starts quickly (fires do not reach large sizes), while those that are far from populated areas remain without any action as they do not pose a great threat. Along roads that are far from villages, fires can reach serious sizes. Additionally, villages in this region are usually not surrounded by forests, but by meadows called "alas", with lower susceptibility to ignition.

Human activity has a greater influence on the fire regimes and differs in the macro- and microscale. The results showed that distance from settlements ($R = -0.97$, and $R = 0.86$), distance from rivers ($R = 0.91$, and $R = -0.54$), and distance from roads ($R = -0.80$, and $R = -0.82$) had a great importance in fire risk assessment. Ajin et al. [43] showed a high correlation between fire occurrence and distance from roads and settlements. The study by Giglio showed that in residential areas and near roads, more human activities are witnessed, and human activity is the most significant factor in the fire outbreak [44]. Studies of the territory of Yakutia in residential areas and near roads show that more human activities are witnessed, and the human activity is the most significant factor in fire outbreaks.

Very large differences were observed between the results of the research on the macro- and microscale. Often the results were quite the opposite. It is probably related to the specificity of the

Yakutia region. The Nyurbinsky region is quite highly urbanized compared to other regions of Yakutia, which may affect the results. Additionally, the regional climatic conditions differed from the average for the entire region. The Nyurbinsky region is characterized by flat terrain, and most of Yakutia is mountainous, which can also affect the conditions for fires. To analyze the macroscale, it is necessary to take into account a wide variety of conditions: climatic, geographical, and human influence. However, on a microscale, these conditions are more homogeneous.

The analysis of the different methods of wildfire risk assessment allows for the identification of the high danger areas in Nyurbinsky and the entire Yakutia. The maximum accuracy was demonstrated by the random forest method with 11 predictors ($R = 0.98$) for the entire territory of Yakutia. The analysis showed that the random forest method gave more accurate results and a much narrower area of the possibility of fires than the MaxENT method, which allows us to propose that this model is preferable. A model created using the maximum entropy method has a small differentiation into zones, which does not allow for its use in practice as uninformative. Parisien used the MaxENT and random forest model to predict fire ignition across the USA [28] with satisfactory results. The MaxENT model was used to assess fire risk in India's Ghats Mountains [32]. In their studies, the authors showed a good correlation between the risk models obtained when using machine learning and forest fires. Due to their power, versatility, and ease of use, random forests are quickly becoming one of the most popular machine learning methods [45]. The studies confirmed that the use of GIS and remote sensing technologies can be used to assess the long-term risk and probability of fires in Yakutia. The results showed that reducing the number of factors during modeling did not show a significant impact on the results in both methods. In the few last years, an increase in the number and severity of fires has been observed as has more years with extreme fire seasons [9]. This situation may also be affected by climate change. On the other hand, a larger number of fires may have an impact on the climate, permafrost ecosystems, and the functioning of indigenous people. In connection with this situation, it is necessary to work on new fire risk assessment methods in the Arctic and subarctic zone of Yakutia. Methods should use global databases of fire, climate, and human activities. For building new models that include all these factors, it is necessary to use different types of datasets, like OSM data, climatic datasets, and remote sensing data from different sensors (MODIS, Landsat, Sentinel). According to Gigović et al., the random forest model could be used at the regional level for forest fire susceptibility mapping [47]. Banerjee et al. claim that the MaxENT method can be used as a decision support tool for stakeholders of forest resources [49]. Pham et al. claim that models that take into account climate variables, vegetation, and human influences can explain fire risk better than those that only account for some of these factors [51].

The objective of work on modeling long-term wildfire risk in the Nyurbinsky region of Yakutia was accomplished. The accuracy of the results depends on the data used, and here available. The results were valid within the limits caused by the data used, the approach was chosen as was the research objective. Despite these limitations, it was possible to obtain interesting results that enabled us to answer the main questions posed by the problem.

5. Conclusions

The problem of the forest fires in Yakutia is not as well studied as in other countries. The results of the research have shown a strong human influence on the risk in this region despite the low population density. Anthropogenic factors showed a high correlation with the occurrence of wildfires

more than climatic or topographical factors. Other factors affect the risk of fires at the macroscale, and others the microscale, which should be considered when modeling.

The random forest method showed better results in the macroscale, however, the maximum entropy model was better in the microscale. The exclusion of variables that do not show high correlation does not always improve the modeling results. The random forest presence prediction model is a more accurate method and significantly reduces the risk territory. The reverse is the method of maximum entropy, which is not as accurate and classifies very large areas as endangered.

Further study of this topic requires a clearer and conceptually developed approach to the application of remote sensing data. Therefore, this work makes sense to lay the foundations for the future of a completely automated fire risk assessment application in the Republic of Sakha. The results can be used for fire prophylactics and planning fire prevention. In the future, to determine the risk well, it is necessary to combine the obtained maps with the seasonal risk determined using indices (for example, the Nesterov index 1949) and the periodic dynamics of forest fires, which Isaev and Utkin studied in 1963. Such actions can help build an application with which it will be possible to determine the risk of wildfire and the spreading of fire during extreme events.

Author Contributions: S.G., and P.J. participated in the conceptualization of the manuscript. S.G. and P.J. participated in the methodology development and software development. P.J. conceived the designed the experiments. S.G. and P.J. wrote the manuscript. S.G. provided supervision on the design and implementation of the research. All authors contributed to the review and improvement of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the French ANR PUR (Polar Urban Centers) in the Work Package 4 focusing on the question of settlements and urban centers risk exposure modeling. The present research received support from Campus France and the CNRS PEPS RICOCHET.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Kasischke, E.S.; Stocks, B.J. *Fire, Climate Change, and Carbon Cycling in the Boreal Forest*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2012; pp.203–204; ISBN 978-0-387-21629-4.
2. Troeva, E.I.; Isaev, A.P.; Cherosov, M.M.; Karpov, N.S. *The Far North: Plant Biodiversity and Ecology of Yakutia*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2010; pp. 203–204; ISBN 978-90-481-3774-9.
3. San-Miguel-Ayanz, J.; Carlson, J.D.; Alexander, M.; Tolhurst, K.; Morgan, G.; Sneeuwjagt, R.; Dudley, M. Current Methods to Assess Fire Danger Potential. In *Wildland Fire Danger Estimation and Mapping*; Series in Remote Sensing; WORLD SCIENTIFIC: 2003; Volume 4, pp. 21–61; ISBN 978-981-238-569-7.
4. Жданко, В.А.; Гриценко, М.В. Метод анализа лесопожарных сезонов: Практические рекомендации. *Л.: ЛНИИЛХ* 1980. (In Russian).
5. Tedim, F.; Leone, V.; Amraoui, M.; Bouillon, C.; Coughlan, M.R.; Delogu, G.M.; Fernandes, P.M.; Ferreira, C.; McCaffrey, S.; McGee, T.K.; et al. Defining Extreme Wildfire Events: Difficulties, Challenges, and Impacts. *Fire* 2018, 1, 9.
6. Oliveira, S.; Oehler, F.; San-Miguel-Ayanz, J.; Camia, A.; Pereira, J.M.C. Modeling spatial patterns of fire occurrence in Mediterranean Europe using Multiple Regression and Random Forest. *Forest Ecol. Manag.* 2012, 275, 117–129.
7. Cardille, J.A.; Ventura, S.J.; Monica, G.T. Environmental and social factors influencing wildfires in the Upper Midwest, United States. *Ecol. Appl.* 2001, 11, 111–127.



8. González-Cabán, A. *Proceedings of the Second International Symposium on Fire Economics, Planning, and Policy: A Global View*; U.S. Department of Agriculture, Forest Service, Pacific Southwest Research Station: Albany, CA, USA, 2008; 720p.
9. Isaev, A.P.; Protopopov, A.V.; Protopopova, V.V.; Egorova, A.A.; Timofeyev, P.A.; Nikolaev, A.N.; Shurduk, I.F.; Lytkina, L.P.; Ermakov, N.B.; Nikitina, N.V.; et al. Vegetation of Yakutia: Elements of Ecology and Plant Sociology. In *The Far North: Plant Biodiversity and Ecology of Yakutia*; Troeva, E.I., Isaev, A.P., Cherosov, M.M., Karpov, N.S., Eds.; Plant and Vegetation; Springer: Dordrecht, The Netherlands, 2010; pp. 143–260; ISBN 978-90-481-3774-9.
10. Cherosov, M.M.; Isaev, A.P.; Mironova, S.I.; Lytkina, L.P.; Gavriilyeva, L.D.; Sofronov, R.R.; Arzhakova, A.P.; Barashkova, N.V.; Ivanov, I.A.; Shurduk, I.F.; et al. Vegetation and Human Activity. In *The Far North: Plant Biodiversity and Ecology of Yakutia*; Troeva, E.I., Isaev, A.P., Cherosov, M.M., Karpov, N.S., Eds.; Plant and Vegetation; Springer: Dordrecht, The Netherlands, 2010; pp. 261–295; ISBN 978-90-481-3774-9.
11. Peters, M.P.; Iverson, L.R.; Matthews, S.N.; Prasad, A.M. Wildfire hazard mapping: Exploring site conditions in eastern US wildland–urban interfaces. *Int. J. Wildland Fire* **2013**, *22*, 567–578.
12. Oliveira, S.; Oehler, F.; San-Miguel-Ayanz, J.; Camia, A.; Pereira, J.M.C. Modeling spatial patterns of fire occurrence in Mediterranean Europe using Multiple Regression and Random Forest. *Forest Ecol. Manag.* **2012**, *275*, 117–129.
13. Phillips, S.J.; Dudík, M. Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. *Ecography* **2008**, *31*, 161–175.
14. Gull, S.F.; Skilling, J. Maximum entropy method in image processing. In *IEE Proceedings F (Communications, Radar and Signal Processing)*; IET Digital Library: London, UK, 1984; Volume 131, pp. 646–659.
15. Атлас сельского хозяйства ЯАССР, А. под редакцией АГ Гущиной. М. Главное Управление Геодезии и Картографии при Совете Министров СССР **1989**. (In Russian).
16. Воробьев К.А.; Ким А.Ю. *Географический Атлас Республика Саха (Якутия)*; Роскартография, 2000. (In Russian).
17. Elovskaiã, L.G. *Классификация и диагностика мерзлотных почв Якутии*; Якутский филиал СО АН СССР: 1987. (In Russian).
18. Giglio, L. MODIS Collection 6 Active Fire Product User's Guide Revision A. Available online: /paper/MODIS-Collection-6-Active-Fire-Product-User%27s-Guide-Giglio/4aaca34ad3bcd557591067399ebc38580eb8286 (accessed on 14 September 2020).
19. Prichard, S.J.; Kennedy, M.C. Fuel treatments and landform modify landscape patterns of burn severity in an extreme fire event. *Ecol. Appl.* **2014**, *24*, 571–590.
20. Martínez, J.; Vega-García, C.; Chuvieco, E. Human-caused wildfire risk rating for prevention planning in Spain. *J. Environ. Manag.* **2009**, *90*, 1241–1252.
21. Syphard, A.D.; Radeloff, V.C.; Keuler, N.S.; Taylor, R.S.; Hawbaker, T.J.; Stewart, S.I.; Clayton, M.K. Predicting spatial patterns of fire on a southern California landscape. *Int. J. Wildland Fire* **2008**, *17*, 602–613.
22. Kwak, H.; Lee, W.-K.; Saborowski, J.; Lee, S.-Y.; Won, M.-S.; Koo, K.-S.; Lee, M.-B.; Kim, S.-N. Estimating the spatial pattern of human-caused forest fires using a generalized linear mixed model with spatial autocorrelation in South Korea. *Int. J. Geogr. Inf. Sci.* **2012**, *26*, 1589–1602.
23. Fick, S.E.; Hijmans, R.J. WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **2017**, *37*, 4302–4315.
24. Illera, P.; Fernández, A.; Delgado, J.A. Temporal evolution of the NDVI as an indicator of forest fire danger. *Int. J. Remote Sens.* **1996**, *17*, 1093–1105.
25. Goward, C.R.; Murphy, J.P.; Atkinson, T.; Barstow, D.A. Expression and purification of a truncated recombinant streptococcal protein G. *Biochem. J.* **1990**, *267*, 171–177.



26. Danielson, J.J.; Gesch, D.B. *Global Multi-Resolution Terrain Elevation Data 2010 (GMTED2010)*; US Department of the Interior, US Geological Survey: Newton, MA, USA, 2011.
27. Jenks, G.F. The Data Model Concept in Statistical Mapping. *Int. Yearb. Cartogr.* **1967**, *7*, 186–190.
28. Parisien, M.-A.; Snetsinger, S.; Greenberg, J.A.; Nelson, C.R.; Schoennagel, T.; Dobrowski, S.Z.; Moritz, M.A. Spatial variability in wildfire probability across the western United States. *Int. J. Wildland Fire* **2012**, *21*, 313–327.
29. Harte, J.; Newman, E.A. Maximum information entropy: A foundation for ecological theory. *Trends Ecol. Evol.* **2014**, *29*, 384–389.
30. Elith, J.; Graham, C.H.; Anderson, R.P.; Dudík, M.; Ferrier, S.; Guisan, A.; Hijmans, R.J.; Huettmann, F.; Leathwick, J.R.; Lehmann, A. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **2006**, *29*, 129–151.
31. Franklin, J. *Mapping Species Distributions: Spatial Inference and Prediction*; Cambridge University Press: Cambridge, UK, 2010; ISBN 978-1-139-48529-6.
32. Renard, Q.; Péliissier, R.; Ramesh, B.R.; Kodandapani, N. Environmental susceptibility model for predicting forest fire occurrence in the Western Ghats of India. *Int. J. Wildland Fire* **2012**, *21*, 368–379.
33. Cutler, D.R.; Edwards T.C., Jr.; Beard, K.H.; Cutler, A.; Hess, K.T.; Gibson, J.; Lawler, J.J. Random forests for classification in ecology. *Ecology* **2007**, *88*, 2783–2792.
34. Курбатский, Н.П. Техника и тактика тушения лесных пожаров. М. Гослесбуиздат **1962**. (In Russian).
35. Tehrany, M.S.; Jones, S.; Shabani, F.; Martínez-Álvarez, F.; Bui, D.T. A novel ensemble modeling approach for the spatial prediction of tropical forest fire susceptibility using logitboost machine learning classifier and multi-source geospatial data. *Theor. Appl. Climatol.* **2019**, *137*, 637–653.
36. Koetz, B.; Morsdorf, F.; van der Linden, S.; Curt, T.; Allgöwer, B. Multi-Source land cover classification for forest fire management based on imaging spectrometry and LiDAR data. *Forest Ecol. Manag.* **2008**, *256*, 263–327.
37. Gai, C.; Weng, W.; Yuan, H. GIS-Based forest fire risk assessment and mapping. In Proceedings of the 2011 Fourth International Joint Conference on Computational Sciences and Optimization; Kunming/Lijiang, China, 5–19 April 2011; pp. 1240–1244.
38. You, W.; Lin, L.; Wu, L.; Ji, Z.; Yu, J.; Zhu, J.; Fan, Y.; He, D. Geographical information system-based forest fire risk assessment integrating national forest inventory data and analysis of its spatiotemporal variability. *Ecol. Indic.* **2017**, *77*, 176–184.
39. Goldarag, Y.J.; Mohammadzadeh, A.; Ardakani, A.S. Fire risk assessment using neural network and logistic regression. *J. Indian Soc. Remote Sens.* **2016**, *44*, 885–894.
40. Andreev, D.V.; Makarova, M.E. GIS technologies application in useful fossils search in the territory of the Republic of Sakha (Yakutia). In Proceedings of the IOP Conference Series: Earth and Environmental Science, Changsha, China, 18–20 September 2020; IOP Publishing: Bristol, UK, 2020; Volume 548, p. 032006.
41. Blanchi, R.; Jappiot, M.; Alexandrian, D. Forest fire risk assessment and cartography. A methodological approach. In Proceedings of the IV International Conference on Forest Fire Research, Luso, Portugal, 18–23 November 2002.
42. Boike, J.; Grau, T.; Heim, B.; Günther, F.; Langer, M.; Muster, S.; Gouttevin, I.; Lange, S. Satellite-derived changes in the permafrost landscape of central Yakutia, 2000–2011: Wetting, drying, and fires. *Glob. Planet. Chang.* **2016**, *139*, 116–127.
43. Ajin, R.S.; Loghin, A.-M.; Jacob, M.K.; Vinod, P.G.; Krishnamurthy, R.R. The risk assessment study of potential forest fire in Idukki Wildlife Sanctuary using RS and GIS techniques. *Int. J. Adv. Earth Sci. Eng.* **2016**, *5*, 308–318.

44. Werf, G.R.; Randerson, J.T.; Giglio, L.; Collatz, G.J.; Mu, M.; Kasibhatla, P.S.; Morton, D.C.; DeFries, R.S.; Jin, Y.; van Leeuwen, T.T. Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009). *Atmos. Chem. Phys.* **2010**, *10*, 11707–11735.
45. Dennison, P.E.; Roberts, D.A.; Reggelbrugge, J.C. Characterizing chaparral fuels using combined hyperspectral and synthetic aperture radar data. In *Proceedings of the 9th AVIRIS Earth Science Workshop, Pasadena, CA, USA, 23–25 February 2000*; Jet Propulsion Lab: Pasadena, CA, USA, 2000; Volume 6, pp. 23–25.
46. Onigemo, A.E.; Santos, S.A.; Pellegrin, L.A.; Abreu, U.G.P.; Silva, E.T.J.; Soriano, B.M.A.; Crispim, S.M.A. Application of vegetation index to assess fire risk in open grasslands with predominance of cespituous grasses in the Nhecolândia sub-region of the Pantanal. *Simpósio Bras. Sens. Remoto* **2007**, *13*, 4493–4500.
47. Gigović, L.; Pourghasemi, H.R.; Drobnjak, S.; Bai, S. Testing a new ensemble model based on SVM and random forest in forest fire susceptibility assessment and its mapping in Serbia's Tara National Park. *Forests* **2019**, *10*, 408.
48. Pourghasemi, H.R.; Kariminejad, N.; Amiri, M.; Edalat, M.; Zarafshar, M.; Blaschke, T.; Cerda, A. Assessing and mapping multi-hazard risk susceptibility using a machine learning technique. *Sci. Rep.* **2020**, *10*, 1–11.
49. Banerjee, P. Maximum entropy method-based forest fire prediction mapping of Sikkim Himalaya. **2020**.
50. Lim, C.-H.; Kim, Y.S.; Won, M.; Kim, S.J.; Lee, W.-K. Can satellite-based data substitute for surveyed data to predict the spatial probability of forest fire? A geostatistical approach to forest fire in the Republic of Korea. *Geomat. Nat. Hazards Risk* **2019**, *10*, 719–739.
51. Pham, B.T.; Jaafari, A.; Avand, M.; Al-Ansari, N.; Dinh Du, T.; Yen, H.P.H.; Phong, T.V.; Nguyen, D.H.; Le, H.V.; Mafi-Gholami, D. Performance evaluation of machine learning methods for forest fire modeling and prediction. *Symmetry* **2020**, *12*, 1022.
52. Ghorbanzadeh, O.; Blaschke, T.; Gholamnia, K.; Aryal, J. Forest fire susceptibility and risk mapping using social/infrastructural vulnerability and environmental variables. *Fire* **2019**, *2*, 50.
53. Pew, K.L.; Larsen, C.P.S. GIS analysis of spatial and temporal patterns of human-caused wildfires in the temperate rain forest of Vancouver Island, Canada. *Forest Ecol. Manag.* **2001**, *140*, 1–18.
54. Zhang, Z.X.; Zhang, H.Y.; Zhou, D.W. Using GIS spatial analysis and logistic regression to predict the probabilities of human-caused grassland fires. *J. Arid Environ.* **2010**, *74*, 386–393.

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).



Annexe 2 : questionnaire en ligne

<https://forms.gle/hfgaTcfFaamEWQTj8>

Новая форма

Благодарим Вас за согласие участвовать в научном исследовании, целью которого является исследование особенностей репрезентации географического пространства в рамках международного научно-исследовательского проекта RICOCHET (В поисках интеграции знаний при наблюдении за изменениями окружающей среды) Франция

Le nom et la photo associés à votre compte Google seront enregistrés lorsque vous importerez des fichiers et que vous enverrez ce formulaire.

Vous n'êtes pas le propriétaire de l'adresse e-mail sebastien.gadal@gmail.com ? [Changer de compte](#)

1. Как называется Ваш населенный пункт?

2. Что означает название Вашего населенного пункта?

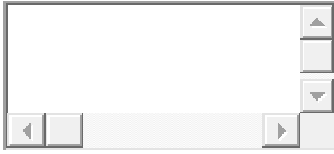
3. Имеется ли у него второе название?

4. Что означает название Вашего надела?

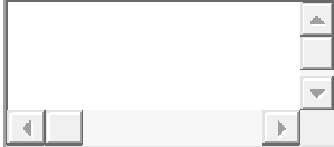
5. Как называются ближайшие реки, речек, ручьев, в бассейне которых Вы живете? Есть ли у них другие названия? Укажите оба названия.

6. Укажите происхождение и значение названий водоемов:

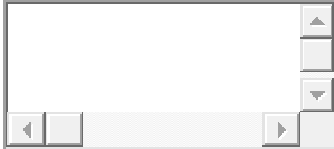
7. Какие слова Вы используете для обозначения разных водных объектов?



8. Как называются ближайшие орографические объекты? Укажите местные термины для определения: Горы Возвышенности (сопки и др.) Низменности Равнины Овраги Обрывы Булгуннях Алаас Острова Полуострова Другие формы рельефа



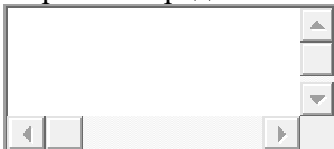
9. Как называются части горы, возвышенности? Укажите местные термины для определения: Подножье Склон Вершина Вулкан Отдельно стоящие останцы Пещеры Гроты Прижимы Ущелья Булгуннях Другие части



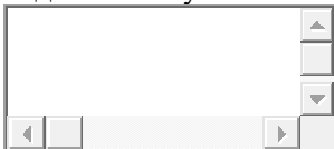
10. Как называются части низменности, равнины? Укажите местные термины для определения: Низменность Равнина Окраина Середина Высокая часть Низкая часть Затопливаемая часть Острова Полуострова



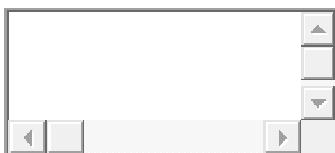
11. Как называются части алааса? Укажите местные термины для определения: Окраина Середина Высокая часть Низкая часть Озеро



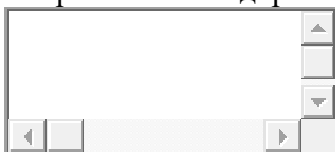
12. Как называются ближайшие к Вашему населенному пункту лесные массивы (топонимы)? Укажите местные термины для определения: Хвойный лес Лиственный лес Смешанный лес Поросль (молодой лес) Горелый лес Вырубки Редколесье Густой лес



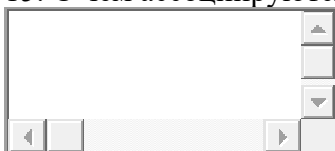
13. Как Вы называете части острова? Укажите местные термины для определения: Начало, Конец, Середина, Озеро внутри острова, Затопливаемая часть, Часть острова, превращающаяся в отдельный остров во время высокой воды



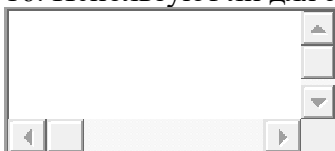
14. Как Вы называете разные виды путей сообщения? Основная дорога
Второстепенные дороги



15. С чем ассоциируются у Вас части света?



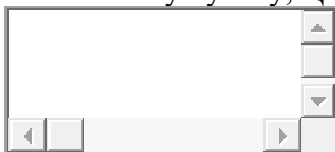
16. Используют ли для этого названия соседних территорий?



17. Какой объект служит для Вас естественным и искусственным ориентиром в определении направления? (Направление ориентировки (от реки, от горы, от леса, дорога, ведущая в населенный пункт и др., а также части света))

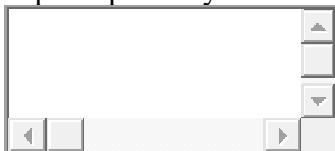


18. Опишите расположение Вашего населенного пункта по отношению к: Соседнему населенному пункту, Центру наслега, Центру района, Города



19. Сделайте схематичный рисунок местности с нанесением основных ориентиров, ландшафтов и указанных вами топонимов, можете нарисовать на листе бумаги сфотографировать и приложить к анкете

20. Какие ландшафты Вы используете для хозяйственной деятельности? Дайте характеристику местности



21. Какие в этой местности есть места сакральные места? Укажите название и примерное местоположение. Это могут быть: Стоянки, захоронения, Название и местоположение, Почитаемые природные объекты, Название и местоположение.

22. Почему эти места являются почитаемыми? Какие суеверия связаны с этими местами?

23. Есть ли в Вашей местности запретные места? Почему они запретные? Укажите название и примерное местоположение .

24. Где у Вас проходят празднования и обряды? Название мероприятия Примерная дата, Местность проведения мероприятия, Характеристика мероприятия

Укажите любым доступным способом на фрагменте космического снимка, указанные вами названия географических объектов(топонимов) и укажите места, подвергаемые природным бедствиям (наводнения и лесные пожары)



N'envoyez jamais de mots de passe via Google Forms.






Ce contenu n'est ni rédigé, ni cautionné par Google. [Signaler un cas d'utilisation abusive](#) - [Conditions d'utilisation](#) - [Règles de confidentialité](#)

Annexe 3


Sébastien Gadal, Moisei Zakharov, Jūratė Kamičaitytė, Yuri Danilov. **Alas Landscape Modeling by Remote Sensing Image Analysis and Geographic Ontology. Study case of Central Yakutia (Russia).** *6th International Conference on Geographic Information Systems Theory, Applications and Management*, Polytechnic Institute of Setúbal/IPS; Knowledge Systems Institute; ATHENA Research & Innovation Information Technologies, May 2020, Online Streaming, Portugal. pp.112-118, ([10.5220/0009569101120118](https://doi.org/10.5220/0009569101120118)).


Alas Landscape Modeling by Remote Sensing Image Analysis and Geographic Ontology. Study case of Central Yakutia (Russia)


Sébastien Gadal ^{1,2}^a, Moisei Zakharov ^{1,2}^b, Jūratė Kamičaitytė ³^c, and Yuri Danilov ²^d
¹ Aix-Marseille Univ, CNRS, ESPACE UMR 7300, Univ Nice Sophia Antipolis, Avignon Univ, 13545 Aix-en-Provence.
² North-Eastern Federal University, 670000 Yakutsk, Republic of Sakha, Russian Federation.
³ Kaunas University of Technology, Kaunas, Lithuania


Keywords: Geographic Ontology, Image Analysis, Knowledge Database, Image Processing, Alas Landscape, Remote Sensing

Abstract: Approaches of geographic ontologies can help to overcome the problems of ambiguity and uncertainty of remote sensing data analysis for modeling the landscapes as a multidimensional geographic object of research. Image analysis based on the geographic ontologies allows to recognize the elementary characteristics of the alas landscapes and their complexity. The methodology developed includes three levels of geographic object recognition: (1) the landscape land cover classification using Support Vector Machine (SVM) and Spectral Angle Mapper (SAM) classifiers; (2) the object-based image analysis (OBIA) used for the identification of alas landscape objects according to their morphologic structures using the Decision Tree Learning algorithm; (3) alas landscape's identification and categorisation integrating vegetation objects, territorial organisations, and human cognitive knowledge reflected on the geo-linguistic object-oriented database made in Central Yakutia. The result gives an ontology-based alas landscape model as a system of geographic objects (forests, grasslands, arable lands, termokarst lakes, rural areas, farms, repartition of built-up areas, etc.) developed under conditions of permafrost and with a high sensitivity to the climate change and its local variabilities. The proposed approach provides a multidimensional reliable recognition of alas landscape objects by remote sensing images analysis integrating human semantic knowledge model of Central Yakutia in the subarctic Siberia. This model requires to conduct a multitemporal dynamic analysis for the sustainability assessment and land management.

^a <https://orcid.org/0000-0002-6472-9955>

^b <https://orcid.org/0000-0002-8916-2166>

^c <https://orcid.org/0000-0002-8821-5764>

^d <https://orcid.org/0000-0003-4806-4938>

1 INTRODUCTION

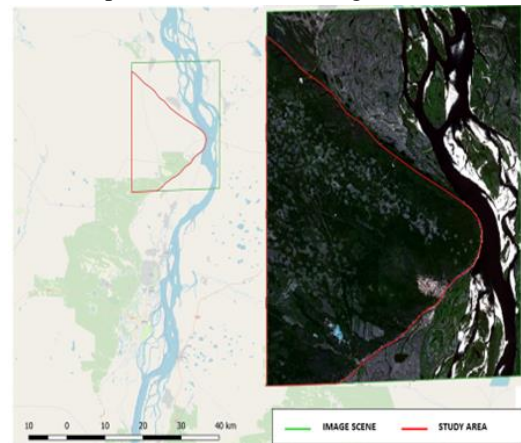
The problem of knowledge integration in remote sensing for the studies of landscape characteristics is one of the key gaps; especially for the modeling of the territorial sustainability assessment and for the ongoing future landscape changes (Konys, 2018). The expert geographic knowledge building based on geographic ontologies is the most suitable approach. It allows the integration of expert-human knowledge for the remote sensing image analysis and modeling of the complexity of the alas landscape system. The approach based geographic ontology can be considered as a specific domain of knowledge in the field of the artificial intelligence. Geographical ontologies used for the alas landscape modeling includes: (a) biophysical geographic components of landscape (e.g., vegetation, moisture, landform (relief), climate conditions, soil (Pashkevich, 2017)); (b) morphologic forms and geometric objects characteristics (e.g., shape, size, textures, topology, geolocation, depression, inclination, slope), and (c) anthropogenic objects related to the land field reality measured (captured) by remote sensing, the human geographic object's identification (semantic meaning) and perceptions. The human geographic object's identification (semantic meaning) was formalized in Russian and Yakutsk languages in a geolinguistic database integrating toponyms, territorial knowledge of local population, census, land uses, maps, etc. (Zamorshikova et al, 2018). This geo-linguistic database including the bunch of heterogeneous knowledge poses the problem of standardization (harmonization). The application in the alas landscape combining the knowledge model, remote sensing and geographic ontologies is based on the theoretical model of ontology (Sinha and Mark, 2010) and on the methodology developed in remote sensing applied to the subarctic urban environment in Yakutsk (Gadal and Ouerghemmi, 2019).

2 ISSUES OF CLIMATE CHANGES OF THE ALAS LANDSCAPE

The interest to study the alas landscape is related to the impacts of climate change and modeling the landscape processes. The alas landscapes are localized in Taiga boreal forest and permafrost territories of the subarctic region of Siberia, Alaska and Canada. They are formed as a result of thawing of the permafrost layer in the exposed parts of the taiga. It makes the alas one of the key indicators of the permafrost degradation and vulnerability of landscapes in the subarctic region (Fedorov, 2019) and has a major contribution on

carbon (CH₄ and CO₂) emission. They have impact on the forms of steppe azonal soils which have a high bio-productivity for the agriculture. Agriculture in the subarctic regions is the main framework of the traditional local lifestyles of Yakutsk and Ekenkii. According to it, the cultural, social and territorial significance of the alas is defined (Desyatkin et al, 2018).

Therefore, the modeling of the alas landscape needs to integrate both the human semantic factors and the geo-physical reality; associating quantitative and semantic parameters defining ontologies of the alas geographic space. The approach for recognizing the geographic ontologies merges manually collected data from the geographic landscape (field observations, population knowledge formalized in geo-linguistic database), maps and remote sensing data: Landsat



series, 5 TM, 7 ETM+, 8 OLI and Sentinel 2. Geographic object ontologies are extracted from the satellite images by semi-automatic or automatic approach of image analysis (Clouard et al, 2013) using machine learning algorithms Spectral Angle Mapper (SAM) on the Landsat series images according to the better performance (overall accuracy). Two approaches were applied for the Sentinel 2 images: (a) the Support Vector Machine (SVM) for the recognition of the geographic objects by pixel-based classification (Li et al 2009, Ouerghemmi et al, 2017) showing the land cover landscape of the alas and the environment around, and (b) the OBIA based on the Decision Tree algorithm (DT) (Vopham et al, 2018) that identified automatically the current different geographic objects defining the actual alas. Results are integrated and merged in GIS. It must be noted that the application of the DT algorithm on the Landsat TM, ETM+, OLI is not relevant because of the too low GSD of 30m.

2.1 Study area

The Central Yakutia lowland is characterized by the widespread distribution of Alas. Throughout the territory there are about 16000 Alases between Vilyuy, Lena and Aldan Rivers. The study area is located at 40

km distance to the North from the city of Yakutsk covering the surface of 30000 hectares (Fig. 1). The study area is a boreal forest represented by pine and larch. The soil basis is permafrost with a layered and lenticular texture, where thermokarst cryogenic processes are developed, which cause the formation of numerous alas.

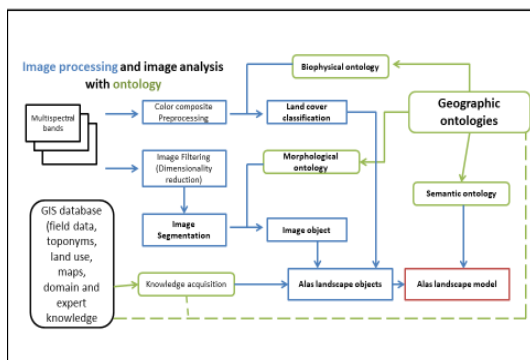
Figure 1. The study area [© Open Street Map]; the extent of the Sentinel 2A RGB scene at 26th of July 2019.

3 IMAGE ANALYSIS OF ALAS LANDSCAPE STRUCTURE

3.1 Methodology

The geographic objects as ontological elements are the basis of the structure and organization of a landscape (Gadal, 2012). In remote sensing they are identified by image processing and image analysis (here by machine learning, landscape object classifications, and OBIA image segmentation) with the knowledge GIS database and the implementation of ontological rules (Fig. 2).

Figure 2. Implementation of image analysis with



geographic ontologies for alas landscape recognition.

The typical alas landscapes consist of the following elementary geographic objects: lake, grassland, woodland, built-up and arable lands for agriculture. First, they are identified and analysed according to land cover classes produced from the biophysical parameters measured spectrally by Landsat series and Sentinel 2 data.

The second level of the geographical objects recognition by object-oriented classification algorithms of DT is resulted from knowledge acquisition (Clouard et al, 2010) structured in the geo-linguistic database and integrated in GIS, and the OBIA to extract the morphology structures of the alas landscape. The categorization of the alas landscapes objects generated is followed-on by the fusion of land cover maps and alas landscape object classification.

The data analysed include the expert and the domain knowledge (landscape theories, concepts, indicative and cognitive analysis products) (Gadal and

Ouerghemmi, 2019). The ontological approach enables us to obtain numerous representations of all aspects of the alas landscape spatial organization through the accurate acquired knowledge structured in the GIS database. It is used both for training, learning and the validation of results, the spatial analysis and modeling in GIS.

The integrity and the logic coherence of all processes taking place in geographic space is ensured by the integration of semantic geographic object ontologies modelled by the geostatistical analysis of geo-linguistics database (toponyms, territorial knowledge of local population, census, land uses, maps, etc.) with the categorization of alas landscape based on GIS knowledge database.

The alas landscape model obtained (Fig. 3) is a multidimensional model including knowledge and geo-linguistic GIS database adapted for machine learning training and simulation of geospatial processes: rural exodus, permafrost degradation, etc. having impact on the landscape's changes and their consequences (sustainability, re-greening).

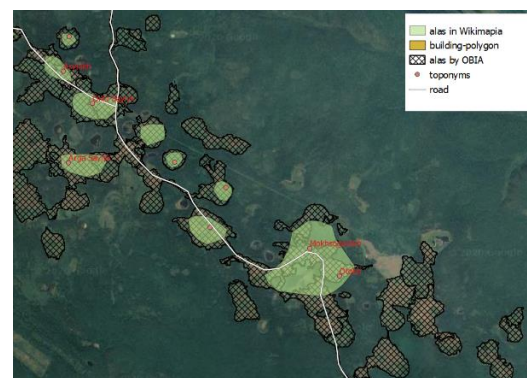


Figure 3. Alas landscape model map with knowledge and geo-linguistic GIS database. Terra-Metrics, DigitalGlobe 2012.

3.2 Image processing

3.2.1 Recognition of objects of landscape land cover change

The recognition of the landscape objects and the change dynamics modeling between 2007 and 2019 is based on the use of the Landsat 5 TM (date: 15.07.2007), Landsat 8 OLI (bands 2 to 7 : VIR-NIR-SWIR) pansharpned at 15m (date: 26.07.2013) and Sentinel 2A (bands 2 to 5 : VIR- NIR at 10m) (date: 24.07.2019) acquired in July.

Landscape land covers produced and mapped by supervised classification are used as training data of the GIS database. Landscape objects are recognized by the spectral signatures related to the geographic objects of the GIS database. Several classifier algorithms were tested: Minimum Distance(MD), Maximum

Likehood(ML), the K-Nearest Neighbors (K-NN), Spectral Angle Mapper (SAM) and Support Vector

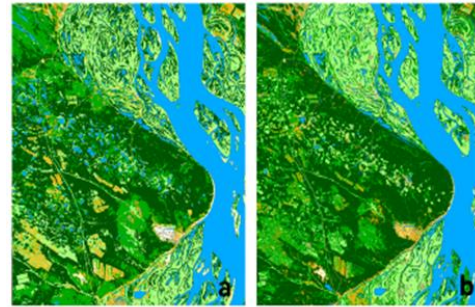


Figure 4. Land cover classification by SAM for Landsat 5 TM in 2007 (a) and Landsat 8 OLI in 2013 (b).

Figure 5. Land cover classification result for Sentinel 2A with SVM classifier.

Machine (SVM).

The results show that the optimal machine learning classifiers are obtained with SAM for the Landsat 5 TM and the Landsat 8 OLI data confirming the results presented by Pertopoulos et al. (2010) (Tab. 1). The most reliable results for the landscape land cover were proceed with the SVM based on Kernel method for the Sentinel 2 images (Pal and Mather, 2005) (Fig x).

Table 1. Results obtained from applied classifiers.

Classification algorithm	O.A. / Kappa		
	2007	2013	2019
MD	72%/0,45	67%/0,44	45%/0.37
ML	43%/0,31	56%/0.39	54/0.36
K-NN	non	non	34%/0.21
SAM	77%/0.57	81%/0.64	67%/0,59
SVM	non	non	83%/0,65

The accuracy of landscape land cover classifications is provided with 50 randomly generated polygons (3x3 pixels size). The random set of vector data (by 3x3 pixels) constitutes the validation data set. They are used for the accuracy assessment of the classification by confusion matrix and Kappa coefficient.

3.2.2 Alas recognition by object-based image analysis

The recognition of alas is based on two criteria that ontologically characterize geographical objects: morphology and their own geometries. They lend themselves to OBIA procedures. The satellite images used are from the VIR and NIR bands of Sentinel 2A, in particular because of their spatial resolution of 10m that enables to recognize the geometric shapes and the structures of the alas. The alas as a unique geographic object is characterized ontologically by its geometry and by the geographic objects that compose it such as lakes, meadows, swamps, the presence of farms and forests surrounding it. These geographic objects are identified by classification (part 3.2.1.) with the landscape cover maps (fig. x, fig. z).

The detection, identification and extraction of alas from their morphological characteristics lends itself well to the OBIA object-oriented machine learning classification algorithms. The OBIA alas recognition consists of four phases: (a) The dimensionality reduction by the Principal Component Analysis (PCA) algorithm, and linear combination of original bands that contains 91-95% of spectral information in first three components (Fig 5a). (b) Smoothing filtering that is applied for the suppression of noise and homogenize statistically the alas regions. This filtering method associates with each pixel of the image the closest local mode in the density distribution of the combined region (Comaniciu and Meer, 1999, 2002). (c) The image segmentation by Large Scale MeanShift non-parametric and iterative clustering method (Michel et al, 2015). The segmentation produces a labeled image with tile-wise processing where neighbor pixels whose range distance is below range radius and optionally spatial distance below spatial radius are merged into the same raster value (Fig.5 c). The average spatial neighborhood radius is 25 meters.

Figure 5. Object recognition from Sentinel 2A (a) RGB of PCA components; (b) MeanShiftSmoothing filtering; (3) Image segmentation of alas.

The most complete extraction of alas is obtained with the following parameters: the spatial radius – 50; the range radius – 25; the Minimum segment size – 10. The main geometrical parameters distinguishing alas and thermokarstic basins concern their shapes: alas has a shape close to the circle, does not have distinct angles, and is not characterized by elongation contrary to thermokarstic shapes. (d) The fourth phase consists in separating the alas from another anthropogenic objects: agricultural parcels. The identification methodology uses a decision tree algorithm. This machine learning algorithm required a dataset that selected 30 polygons generated using the GIS database, more specifically the alas toponyms, topographic maps and Open Street Map data. The classification structure defined by the decision tree is estimated from the training data using a statistical procedure of overall mean and standard deviation for 3 PCA components. The decision rules developed are used to associate each segment of the image with one of the geographic object classes. The O.A. with Sentinel 2A data is =88.4%; a Kappa coefficient is 0,83 for the Decision Tree classifier.

The result is two broad types of geographic objects in which elementary ontological and geographic structures are reflected: alas and arable lands of agriculture. The information was obtained through the interpretation of remote sensing data by artificial intelligence with the integration of expert knowledge for learning and validation of the results obtained. This GIS-structured information set provides the basis for a comprehensive spatial analysis using semantic ontology to model the landscape dynamics of the boreal forests of alas and permafrost.

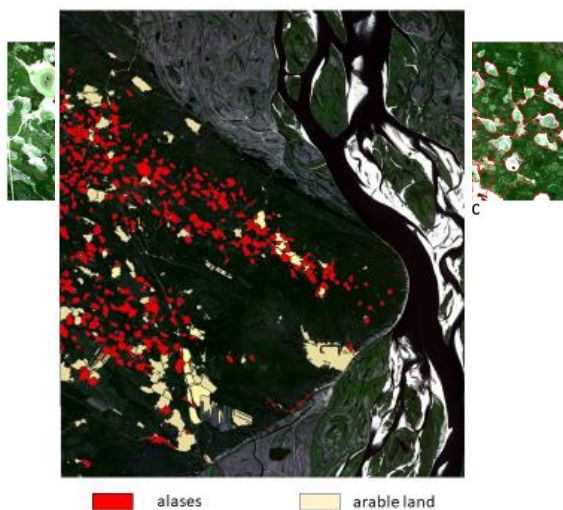


Figure 6. Alas identification by OBIA with Sentinel 2A image.

4 ALAS LANDSCAPE MODELING

In this part, some applications of the image analysis data obtained in the modeling of the alas landscapes are presented. The ecological status of the alas landscapes plays an important socio-economic role, as well as in the dynamic processes of the permafrost. The results of spatial remote sensing image analysis using GIS-integrated geographic ontologies are designed to serve as a basis and environment for this research. Thus, the modeling of geo-linguistic landscape features has been developed in GIS with the creation of an extensive geographically referenced database. It aims to understand better the land use and to measure the complexity of landscape structure.

The first method consists of performing modeling analysis land covering changes within alas. The 4 categories were identified according to the structure of alas vegetation: (a) Termokarst lake, when the alas basin is completely filled with water; (b) Complex alas, when vegetation changes from steppe to coastal-aquatic; (c) herbaceous alas, when vegetation is relatively uniform throughout all alas territories; (d) pasture alas, with active arable or hay areas. The application of knowledge about the vegetation species of alas allows interpreting land cover classes and to assess the stability and transformation of the alas landscape. In total, the alas include 5 typical land cover classes. A highly moistened alas section belongs to the *grass*, less often to the *forest* due to close vegetative indicators (chlorophyll intensity). In some cases, the presence of forests inside the alas indicates the presence of a hill, that are indicators of the presence of permafrost heaving tubercles – pingo. The *arable land* is the middle alas section, the most bio-productive section often used as a forage base for cattle-breeding and horse-breeding. The peripheral portion of alas with a lack of moisture forms the azonal steppe vegetation and stands out as a class of *bare soils*. Table 2 shows examples of alas with the percentage structure of the land cover and the corresponding category.

This categorization allows us to consider the stage of spatio-temporal development of alas, water content, and to obtain data on anthropization and economic activity. This model also has a perspective for the development of simulation strategies for the transformation of plains landscapes of permafrost.

Table 1. Examples of alas landscape categorization by land cover structure.

The second method is based on landscape semantic analysis using the geo-linguistic database. The alas landscapes are the main source of livelihood for the

Name	Year	Land cover classes (%)					Alas category
		forest	Arable land	grass	barren soil	water	
Ton-Bas (Тон-Бас)	2007	17	2	34	0	48	termokarst lake
	2013	18	6	23	17	35	complex alas
	2019	5	10	63	20	2	complex alas
Mokhsogolokh (Моxcoгoллoх)	2007	53	0	9	0	37	termokarst lake
	2013	1	0	99	0	0	herbaceous alas
	2019	5	38	57	0	0	pasture alas
Sier-Bie (Cиep-Бие)	2007	1	0	81	0	18	herbaceous alas
	2013	1	45	53	0	1	pasture alas
	2019	1	47	42	0	11	pasture alas

local population, which implies the presence of a cultural layer of information of the geographical ontology, within the framework of local knowledge about the geographical space. Semantics provides a transition from the object level of the geographical ontology, where the main concept is geographical objects represented in our case as land-use classes - the alas - and the results of their modeling according to the expert knowledge (alas categories, vegetation species), to the spatial level of the geographical ontology. This territorial knowledge generates geographic models whose recognition and identification constitute a new challenge for artificial intelligence in space and airborne remote sensing.

The spatial organization of the local economy involves the location of villages and farms in the alas. Due to their rarity, they cannot be recognized by image processing from medium spatial resolution images such as Sentinel 2A/B, Landsat 5/6/7/8, and are included in the class of arable (agricultural) land. When their presence was known, we reclassified the pixel sets of the arable (agricultural) land value into the buildings class by comparing them with the semantic layer of settlements and sayylyks (summer houses).

An initial geo-linguistic analysis also allows us to identify the alas used in agriculture, which may have been abandoned recently. Compared to satellite images analysis, the absence of a semantic layer can indicate the age and time of alas formation. The GIS semantic database also allows the dissemination of

remote sensing knowledge and facilitates the interpretation of landscape ontology. It provides a semantic description of the results of analysis and processing of remote sensing images.

5 CONCLUSIONS

Geographic ontologies image analysis in artificial intelligence is a necessary part of knowledge that combines multi-level and heterogeneous concepts of an object and a geographic space, and as well geo-linguistics in order to implement a systematic approach to the modeling of alas landscapes. Although, the developed model confronts the technical issues, such as the standardization and harmonization of heterogeneous information flows, it also establishes the rules for solving decision-making problems, for environmental and agricultural monitoring under the conditions of continuously expanding permafrost in subarctic climate.

his research significantly contributes to the development of application of knowledge model to improve remote sensing image processing analysis and landscape modeling for the territorial and geographic studies. One of the main challenges is to hybridize geographical knowledge with artificial intelligence used to analyze the landscapes and the cross processes of anthropization and geophysical evolution of permafrost by spatial and airborne remote sensing.

ACKNOWLEDGEMENTS

This research was supported by the French National Research Agency (ANR) through the PUR project (ANR-14-CE22-0015), the Russian Science Foundation № 15-18-20043 "Landscape Ontology: Semantics, Semiotics and Geographical Modeling", and the Vernadski Grant of the French Embassy in the Russian Federation.

REFERENCES

- Clouard, R., Renouf, A., Revenu, M., 2010. An Ontology-Based Model for Representing Image Processing Objectives. *International Journal of Pattern Recognition and Artificial Intelligence, World Scientific Publishing*, pp.1181-1208.
- Comaniciu, D. and Meer, P. 1999. Mean shift analysis and applications. *International Conference on Computer Vision*, 2, pp. 1197-1203.
- Comaniciu, D., and Meer, P. 2002. Mean shift: A robust approach toward feature space analysis. *IEEE*



- Transactions on pattern analysis and machine intelligence, 24(5), pp. 603-619.
- Desyatkin, A., Takakai, F. and Nikolaeva, M. 2018. Landscape Microzones within Thermokarst Depressions of Central Yakutia under Present Climatic Conditions. *Geosciences 2018*, 8. MDPI
- Fedorov, A., 2019. Permafrost Landscapes: Classification and Mapping. *Geosciences 2019*, 9. MDPI.
- Gadal, S., 2012. Geographic Space Ontology, *Locus Object, and Spatial Data Representation Semantic Theory*. Universal Ontology of Geographic Space: Semantic Enrichment for Spatial Data, IGI Global, pp.28-52, 9786-1-4666-032-27-1.
- Gadal, S., and Ouerghemmi, W., 2019. Knowledge Models and Image Processing Analysis in Remote Sensing: Examples of Yakutsk (Russia) and Kaunas (Lithuania). *In Proceedings of the 5th International Conference on Geographical Information Systems Theory, Applications and Management*, Heraklion, Crete – Greece. pp. 282-288.
- Konys, A., 2018. An Ontology-Based Knowledge Modelling for a Sustainability Assessment Domain. *In Sustainability Journal 10*.
- Kordjamshidi, P., Moens, M-F. 2014. Global Machine Learning for Spatial Ontology Population. *Journal of Web Semantics* pp. 1-29
- Li, J., Carlson, B.E., and Laci A.A., 2009. A study on the temporal and spatial variability of absorbing aerosols using total ozone mapping spectrometer and ozone monitoring instrument aerosol index data. *J. Geophys. Res.*, 114.
- Michel, J., Youssefi, D., & Grizonnet, M. (2015). Stable mean-shift algorithm and its application to the segmentation of arbitrarily large remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 53(2), pp. 952-964.
- Ouerghemmi, W., Gadal, S., Mozgeris, G., Jonikavičius, D., and Weber, C., 2017. Urban objects classification by spectral library: feasibility and applications. In 2017 Joint Urban Remote Sensing Event (JURSE), Dubai, pp. 1- 4.
- Sinha, G., Mark, D., 2010. Toward A Foundational Ontology of the Landscape. MDPI.
- Pal, M. and Mather, P. 2005. Support vector machines for classification in remote sensing. *In International Journal of Remote Sensing Vol.5*. pp. 1007-1011
- Pashkevich, M., 2017. Classification and Environmental Impact of Mine Dumps. *In Assessment, Restoration and Reclamation of Mining Influenced Soils*. pp. 1-32
- Petropoulos, G., Vadrevu, K., Xanthopoulos, G., Karantounias, G., Scholze, M. 2010. A Comparison of Spectral Angle Mapper and Artificial Neural Network Classifiers Combined with Landsat TM Imagery Analysis for Obtaining Burnt Area Mapping. *Sensors*. pp. 1967-1985.
- Puebla-Martinez, M., Perea-Ortega, J., Simon-Cuevas, A., and Romero, F. 2018. Automatic Expansion of Spatial Ontologies for Geographic Information Retrieval. *Communications in Computer and Information Science*. pp. 659-670
- VoPham, T., Hart, J., Laden, F. and Chiang, Y. 2018. Emerging trends in geospatial artificial intelligence (geoAI): potential applications for environmental epidemiology. *Environ Health 17*, 40
- Zamorshchikova, L., Khokholova, I., Ikonnikova, A., Samsonova, M. and Lebedeva, V. 2018. Toponymic Landscape of Central Yakutia: Etymological Analysis of Geographical Names. *Journal of Siberian Federal University. Humanities & Social Sciences*.



Annexe 4 :

Sébastien Gadal, Moisei Zakharov, Jurate Kamicaityte, Antonina Savvinova, and Yuri Danilov, **Environmental Vulnerability Modeling in the Extensively Urbanized Arctic Center Integrating Remote Sensing, Landscape Mapping, and Local Knowledge**, vEGU 2021, 19-30 April 2021, online.



EGU21-16268

EGU General Assembly 2021

© Author(s) 2021. This work is distributed under the Creative Commons Attribution 4.0 License.



Environmental Vulnerability Modeling in the Extensively Urbanized Arctic Center Integrating Remote Sensing, Landscape Mapping, and Local Knowledge

Sébastien Gadal^{1,2}, Moisei Zakharov^{1,2}, Jurate Kamicaityte³, Antonina Savvinova², and Yuri Danilov²

¹Aix-Marseille Université, CNRS ESPACE UMR 7300, France (sebastien.gadal@univ-amu.fr)

²North-Eastern Federal University, 670000 Yakutsk, Republic of Sakha, Russian Federation

³Kaunas University of Technology, Faculty of Civil Engineering and Architecture, 51367 Kaunas, Lithuania

Arctic extensively urbanized centers are subject to the impact of many negative environmental phenomena progressing in terms of global climate change and regional development in Yakutia in the context of poor and missing databases. For this reason, the modeling of the risk exposures is based on combining the remote sensing, and local knowledge of inhabitants. According to the occurrences of the natural hazards, the territorial management and the decision-making system require the identification and assessment of natural risks to which the rural populations localized in the towns and villages are exposed, for example, in the urban center of Khamagatta located at 70km to the North from Yakutsk near the Lena River. The main environmental vulnerability exposures are seasonal: springtime floods between May and June, the forest fires from June to August, the cyclic permafrost degradation, and river erosion impacts.

The current vulnerability impacts, damages to the lands and the settlements, and the populations risk exposures are analyzed from the maps of vulnerabilities created from remote sensing satellite Sentinel 2A/B series, with the local knowledge of the inhabitants of Khamagatta who lived and perceived all events. All the data generated, maps, models of vulnerability exposures, and local knowledge are integrated, combined, and merged into the geographic information system (GIS). The GIS modeling combines the risk of natural hazards and the damages, and the risk knowledge and perceptions of the inhabitants. Land uses, Landscape classification, and the land cover is made by Object-Based Image Analysis (OBIA) using an optical time series of Sentinel 2 images (2015-2020) including the population knowledge for the recognition of the environmental vulnerabilities. The methodological approach included the participation of local people in workshops through discussion and participatory mapping, questionnaires, and interviews in two



stages. The first stage included the development of the knowledge database for a comprehensive understanding of the life of the local population, including the forms of adaptation to the negative natural phenomena. The collected information is delocalized and integrated into the GIS. The second stage consisted of validation and discussion, including stakeholders (municipality and rescue services) to increase the reliability and legitimacy of the research results.

Perceptions of the inhabitants of Khamagatta are correlated with the maps of risk exposures generated by remote sensing to increase the accuracy of the environmental process modeling and landscape classification. The combination of the environmental change dynamics, the impacts on the towns and villages with the human perception and experience constitutes the main base supporting the prevention mapping of the natural hazards. This data could be very useful in planning the development of Arctic towns and villages and proposing evolution scenarios and urban planning models and strategies for increasing their resilience and adaptation to climate change consequences.



Annexe 5 :

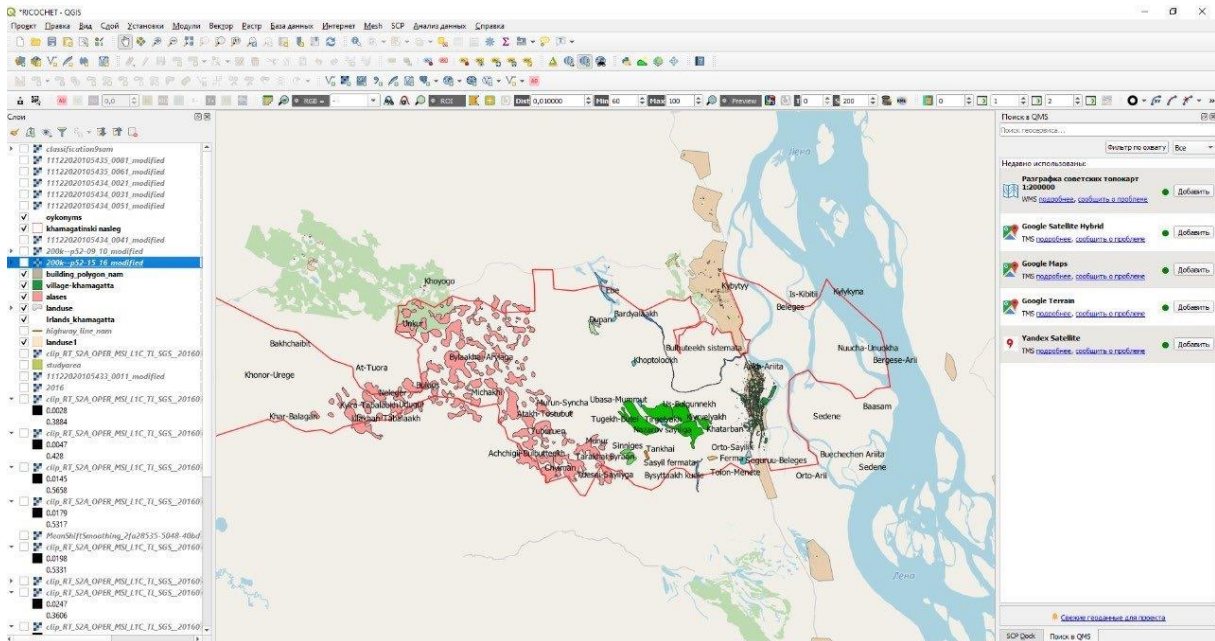
Sébastien Gadal, Moisei Zakharov, Jurate Kamicaityte. Human Perception and Environmental Risk Index Modelling with GIS in Yakutsk Urban Region (North-Eastern Siberia), EUGEO 2021, 26 June-1 July 2021, Prague.

One of the challenges of human security in the Arctic regions, and more specifically in the Eastern Siberian Arctic regions, concerns the assessment of inhabitants' exposure to recurrent environmental risks for the urban population in a geophysical environment characterized by notable change dynamics. The latter is known as permafrost degradation, water flooding (river, geomorphologic canals, and plateau's erosion flooding) and ice breakup, erosion, and forest fires. Urban centers and villages are confronted with a multiplicity of natural hazards. However, the evaluation of their exposure is partial or non-existent; hence there is the need to create indices of potential exposure to environmental risks. They are built by spatial analysis of survey data on the perception and knowledge of the local population and remote sensing that are integrated into Geographic Information Systems. Knowledge and understanding of how local people perceive the natural and anthropogenic environment in the context of climate change and its consequences are some of the key factors in creating effective urban planning strategies for the cities in the high-risk zones of environmental hazards. The urban region of Yakutsk is the test study area for the construction of these indicators of human security in the Siberian Russian Arctic. The large urban area with more than 350,000 people in one of the coldest regions in the world is confronted by the acceleration of global warming and urbanization on a recurrent and seasonal basis. This hazardous occurrence is combined with the risks of forest fires, erosion, floods, and permafrost degradation affecting infrastructures, settlements, industrial and agricultural areas.

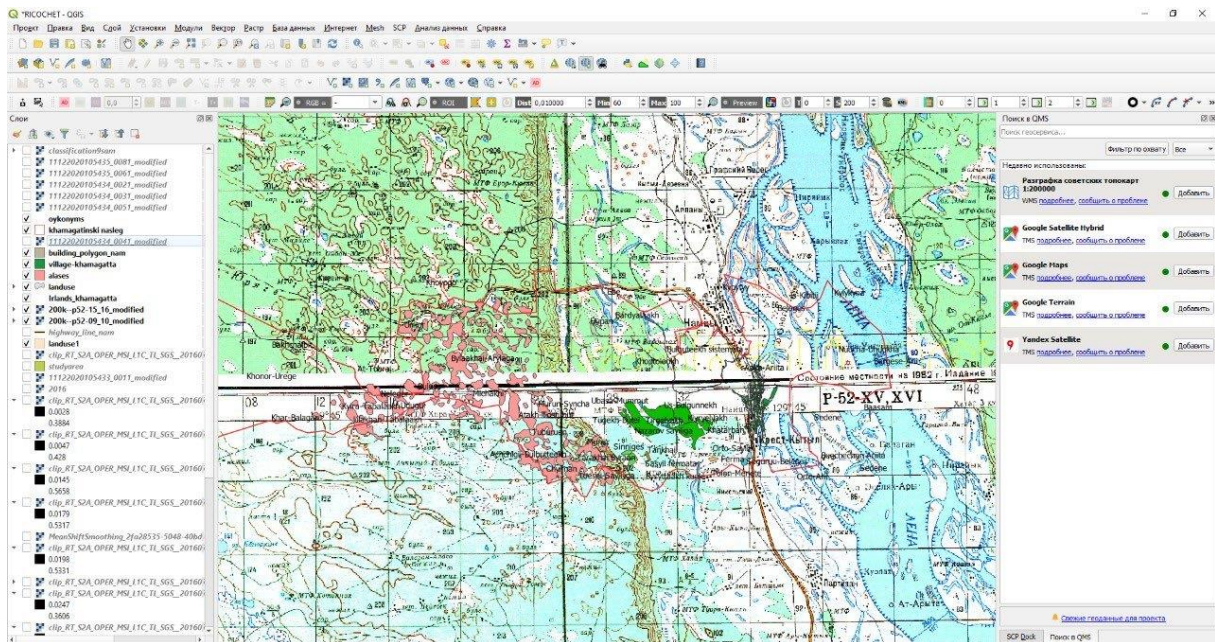
Keywords: Human security environmental risk index, Arctic urban centers, Perceptive knowledge, GIS, Spatial Analysis.



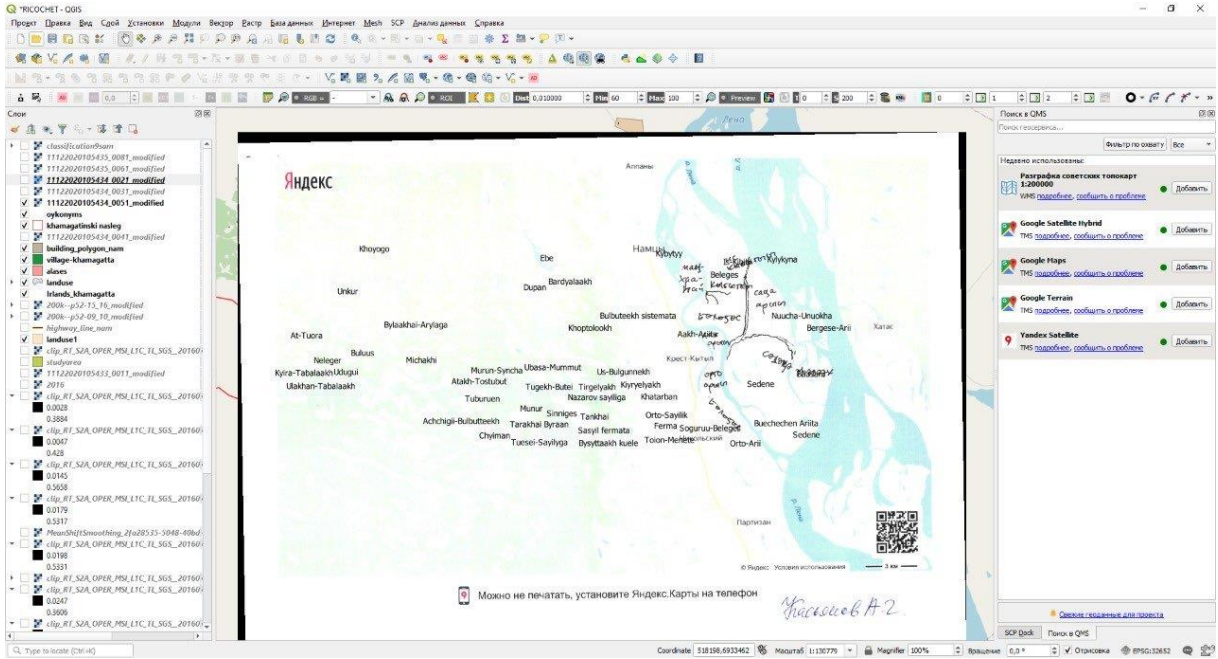
ANNEXE 6 : Extractions du Khamagatta GIS



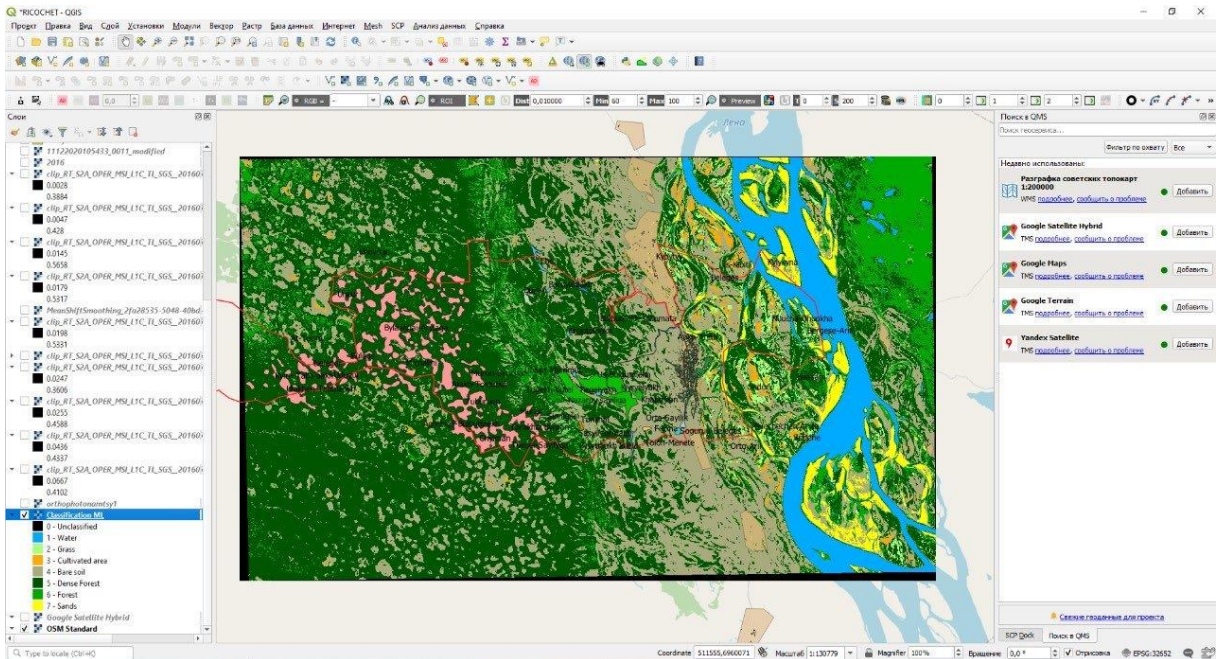
Cartographie d'usage des sols issue de cartographies participatives



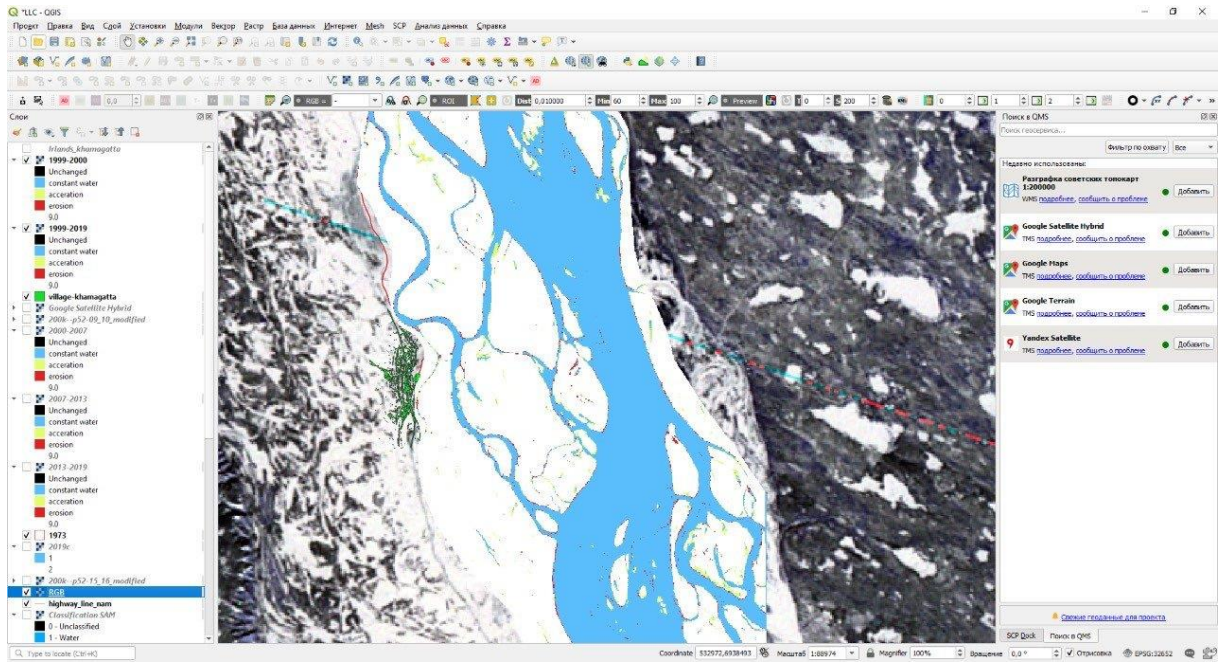
Cartes topographiques soviétiques de 1980



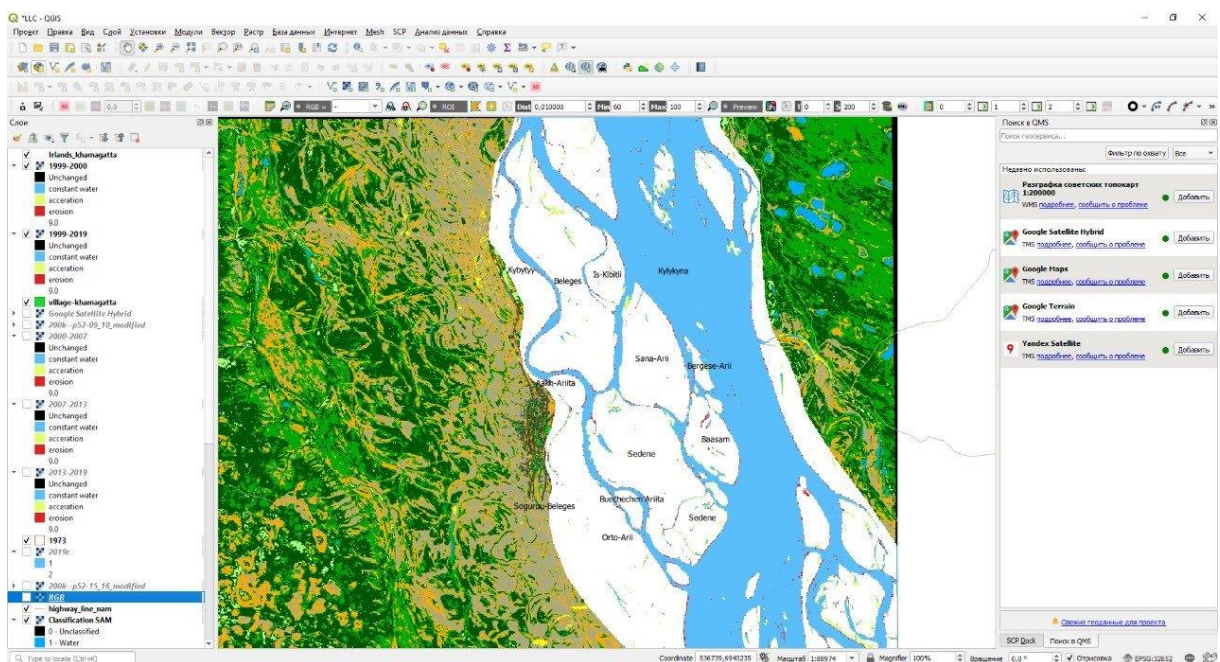
Cartographie vécue d'un des participants



Superposition entre la cartographie d'usage des sols (issue de cartographies participatives), et la spatio-cartographie de l'occupation des sols produites à partie d'images Sentinel 2.



Spatio-cartographie de l'érosion et des accumulations de charges sédimentaires.



Spatio-cartographie de l'occupation des sols et de la spatio-carte mesurant érosion et accumulation de charges sédimentaires.