



**UNIVERSITÉ
DE GENÈVE**

GRID
Geneva



mapx

Automatisation de l'intégration de données dans MapX (une plateforme de données spatiales pour la gestion durable des ressources naturelles)



Certificat Complémentaire en Géomatique
Rapport de stage

Présenté par **Florian Franziskakis**
Superviseur: **Dr. Pierre Lacroix**

Semestre d'automne 2018

RÉSUMÉ

Ce document est un rapport de stage effectué au sein du GRID – Genève dans le cadre du certificat complémentaire en géomatique de l'Université de Genève. Il présente une brève synthèse des différentes missions conduites lors de ce stage en rapport avec la plateforme MapX, une plateforme pour la gestion durable des ressources naturelles. Ces différentes missions ont eu trait à l'automatisation de processus de traitement et d'intégration de données spatiales dans la plateforme, données spatiales en lien avec les ressources naturelles, ou encore avec les domaines environnementaux et humanitaires.

ABSTRACT

This document is the summary and report of an internship within the GRID – Geneva service in the frame of the complementary certificate of geomatics of the University of Geneva. It compiles the various missions conducted during this internship concerning the MapX platform, a cloud solution for mapping and monitoring the sustainable use of natural resources. These different missions were linked to the automatization of data treatment and integration processes into the platform, especially spatial data related to natural resources, as well as environmental and humanitarian fields.

Table des matières:

RÉSUMÉ	2
ABSTRACT	2
1. INTRODUCTION	6
1.1. Le GRID – Genève	6
1.2. La plateforme MapX.....	6
1.2.1. Généralités	6
1.2.2. Visualisation de données.....	7
1.2.3. Interactivité des données	8
1.2.4. Source et intégrité des données.....	10
1.3. Objectifs du stage.....	10
2. PROJETS	12
2.1. Collaboration avec le Geneva Water Hub	12
2.1.1. Présentation	12
2.1.2. Méthode	12
2.1.3. Résultats	13
2.1.4. Bilan	14
2.2. Intégration des données KoBoCollect dans MapX	15
2.2.1. Présentation	15
2.2.2. Méthode	16
2.2.3. Résultats	16
2.2.4. Bilan	16
2.3. Intégration des données FEWS NET dans MapX	17
2.3.1. Présentation	17
2.3.2. Méthode	18
2.3.3. Résultats	19
2.3.4. Bilan	19
2.4. Travail de recherche sur le secteur minier artisanal en RDC	19
2.4.1. Présentation	19
2.4.2. Méthode	20
2.4.3. Résultats	21
2.4.4. Bilan	22
2.5. Tests des calculs de l'aire d'intersection avec l'outil PixOp.....	22
2.5.1. Présentation	22

2.5.2.	Méthode.....	23
2.5.3.	Résultats.....	23
2.5.4.	Bilan.....	24
2.6.	Publication de données dans MapX.....	25
2.6.1.	Présentation.....	25
2.6.2.	Méthode.....	26
2.6.3.	Résultats.....	26
2.6.4.	Bilan.....	27
3.	BILAN GÉNÉRAL DU STAGE.....	28
4.	RÉFÉRENCES.....	29
5.	REMERCIEMENTS.....	30
6.	ANNEXES.....	31

Liste des figures:

Figure 1: Visualisation dans MapX de la couche "Urban Heat Island UHI Summer Daytime Urban Rural Temperature Difference" dans le projet "WESR:Climate".	7
Figure 2: Affichage des informations relatives à un pixel pour la couche "Accumulated Precipitation Estimates (CHIRPS)" dans le projet "World".	8
Figure 3: Exemple d'un dashboard sur la couche "Eastern DRC ASGM sites" dans le projet "Minamata Convention on Mercury".	9
Figure 4: Exemple d'une slide extrait de la story map "Impact of mining activity on the forest cover" dans le projet "UN Environment Democratic Republic of Congo".	9
Figure 5: Vue MapX de la couche des évènements de conflits armés, zoom sur l'Irak, la Syrie et la Palestine.	13
Figure 6: Liste des évènements présentés dans la story map.	14
Figure 7: Exemple issu de la story map concernant un évènement au Yemen en juillet 2018.	14
Figure 8: Processus de traitement des données de KoBo vers MapX.	16
Figure 9: Exemple de la situation pour la région d'Afrique de l'est en juin 2018.	18
Figure 10: : Processus de traitement des données de FEWS NET vers MapX.	19
Figure 11: Rapport IPIS 2015 (IPIS 2016).	20
Figure 12: Capture d'écran de l'interface présentant la carte interactive créée par IPIS. Sur la droite, un menu permet de filtrer les sites miniers artisanaux selon des critères prédéfinis.	21
Figure 13: Vues MapX montrant A) la localisation des conflits (ACLED), B) L'interférence de groupes armés dans les sites miniers artisanaux (IPIS) et C) Les sites miniers artisanaux classés par minéral principal (IPIS).	22
Figure 14: Comparaison du niveau de complexité des géométries de plusieurs couches.	24
Figure 15: Fenêtre de configuration du style d'une vue dans MapX, ici la vue "Area of millet crop (hectares)" dans le projet "Risk - Niger".	25
Figure 16: Comparaison entre les styles de la couche "Area of millet crop (hectares)" définis dans A) QGIS et B) MapX dans le projet "Risk – Niger".	26

1. INTRODUCTION

Ce rapport fait état des missions réalisées pour un stage d'une durée de trois mois et demi, entre début août et mi-novembre 2018, au sein du GRID (Global Resource Information Database) de Genève et plus particulièrement pour la plateforme MapX. Ce stage s'inscrit dans le cadre du [certificat complémentaire en géomatique](#) de l'Université de Genève, formation de 30 ECTS qui comprend 18 ECTS dévolus à des cours et 12 pour lesquels un travail de recherche ou un stage doit être effectué.

1.1. Le GRID - Genève

Le centre du GRID - Genève fait partie de la Division Science du Programme des Nations Unies pour l'environnement (UN Environnement) depuis sa création en 1985. UN Environnement est la branche de l'Organisation des Nations Unies (ONU) qui s'occupe de la gouvernance environnementale au niveau global. Les missions de UN Environnement sont, entre autres, de coordonner les activités de l'ONU pour l'environnement, d'assister les pays à l'établissement de politiques environnementales et du développement durable.

Le GRID – Genève est depuis 1998 un partenariat entre UN Environnement, l'Université de Genève (UNIGE) et l'Office Fédéral de l'Environnement de la Confédération suisse (OFEV). Ses missions sont notamment de fournir des données scientifiques valables et pertinentes pour supporter les décisions en matière d'environnement et de développement durable. Les activités du GRID – Genève varient d'une échelle locale à globale mais sont en premier lieu axées sur les besoins des institutions partenaires.

Pour remplir ces missions, le GRID-Genève utilise des outils d'analyses géospatiales tels que les systèmes d'information géographique (SIG), la télédétection (remote sensing) et l'analyse statistique, tous intégrés dans des infrastructures de données spatiales (SDI) modernes.

1.2. La plateforme MapX

[MapX](#) est une plateforme de données spatiales pour la gestion durable des ressources naturelles développée par le GRID-Genève en partenariat avec UN Environnement et la Banque Mondiale. Cette plateforme en libre accès permet le partage, l'analyse et la visualisation de données spatiales dans un cadre de neutralité garanti par les Nations Unies. Elle permet également de regrouper les informations nécessaires aux parties prenantes dans la prise de décision.

1.2.1. Généralités

Initialement développée pour aider à la gestion et à la transparence de l'utilisation des ressources naturelles, MapX a étendu son champ d'action et s'insère désormais dans un cadre plus large qui concerne le développement durable, la biodiversité, la gestion des risques naturels ainsi que les enjeux environnementaux et humanitaires divers. Actuellement, MapX est engagé dans plus de 12 grands projets, répartis dans 8 pays et totalise environ 130 couches de données.

[L'application MapX](#) permet la visualisation et l'analyse de couches de données géospatiales dans des environnements dédiés (des projets) qui peuvent être des étendues géographiques (Monde, pays), des programmes (UN Biodiversity Lab, Integrated Strategic Environment Assessment project répartis dans de nombreux pays) ou de plus petits projets "ponctuels" tels que le Geneva Water Hub. A travers ses différentes fonctionnalités, MapX apporte des solutions techniques pour favoriser la compréhension de processus souvent complexes, appuyés et étayés par des données fiables.

Les données géospatiales présentes dans MapX peuvent être des couches vectorielles, raster ou encore en code personnalisé (custom code). Des fonctionnalités supplémentaires permettent par ailleurs d'apporter des outils interactifs comme les dashboards (tableaux de bord fixes ou interactifs qui présentent les données sous forme de widgets en parallèle de la carte) ou les story maps (plus complexes sous forme de succession de slides mélangeant des vues, du texte et divers contenus multimédias dans un but de communication).

Les données présentes dans MapX sont affiliées aux projets dans lesquels elles sont pertinentes. La possibilité pour un utilisateur de les visualiser et de les modifier dépend donc des droits qu'il possède dans tel ou tel projet. Plusieurs statuts d'utilisateurs existent (public, membre, éditeur, administrateur) et garantissent une gestion organisée ainsi que la confidentialité parfois nécessaire pour certaines données sensibles.

1.2.2. Visualisation de données

MapX permet la visualisation de données géospatiales dans une interface aisément compréhensible pour un utilisateur possédant des compétences limitées en SIG. Les données sont visualisées grâce à des vues, basées sur des couches sources stockées dans la base de données MapX ou venant de sources externes. Chaque vue possède un style modifiable par les utilisateurs possédant les droits d'édition sur la couche source.

En plus des vues publiques, un utilisateur peut charger ses propres couches de données dans MapX et créer des vues basées sur ces couches. Les formats supportés par MapX sont nombreux (GeoJSON, GeoPackage, ESRI Shapefile, KML, GPX) et le chargement des couches peut être fait rapidement, tout comme la création et la configuration d'une vue.

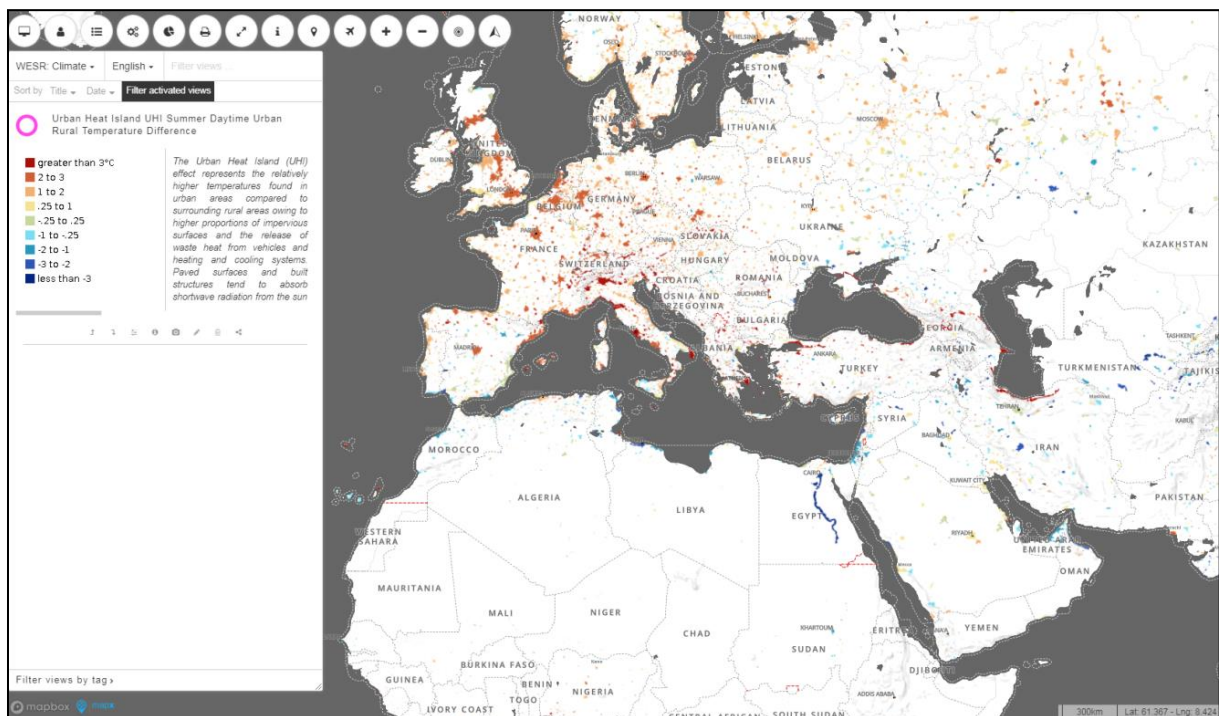


Figure 1: Visualisation dans MapX de la couche "Urban Heat Island UHI Summer Daytime Urban Rural Temperature Difference" dans le projet "WESR:Climate".

1.2.3. Interactivité des données

L'interface MapX présente une barre d'outils (en haut à gauche), une boîte de légende avec la liste des vues (à gauche), une échelle et la position de la souris (en bas à droite). Lorsque l'utilisateur clique sur la carte, les informations relatives aux données s'affichent à l'endroit sélectionné. La barre d'outils permet plusieurs opérations comme l'affichage de la liste des vues, l'affichage de la boîte à outils, l'affichage des dashboards, la capture de l'écran et d'autres fonctionnalités.

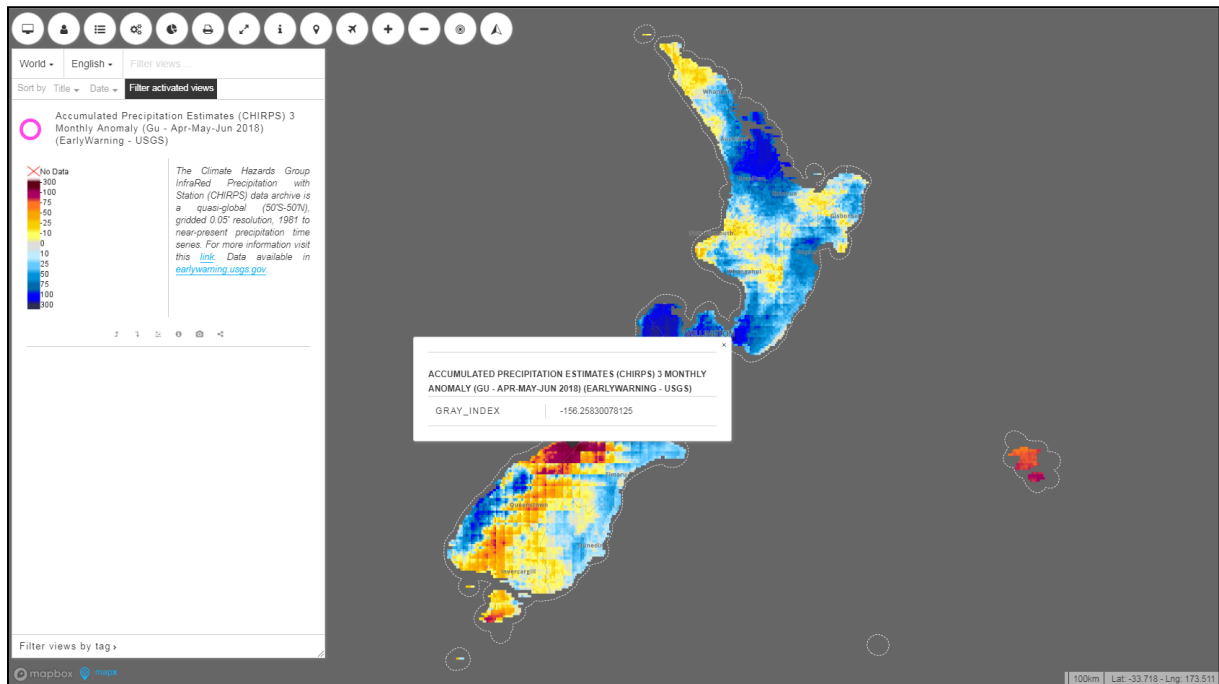


Figure 2: Affichage des informations relatives à un pixel pour la couche "Accumulated Precipitation Estimates (CHIRPS)" dans le projet "World".

Les dashboards permettent d'afficher des informations supplémentaires et ainsi apporter des compléments aux données (notamment sous forme de graphiques, d'images, et d'informations textuelles). Ces dashboards peuvent être dynamiques et s'affichent lorsque l'utilisateur clique sur la carte, ou alors rester fixes. La création de dashboards nécessite des connaissances en programmation (JavaScript).

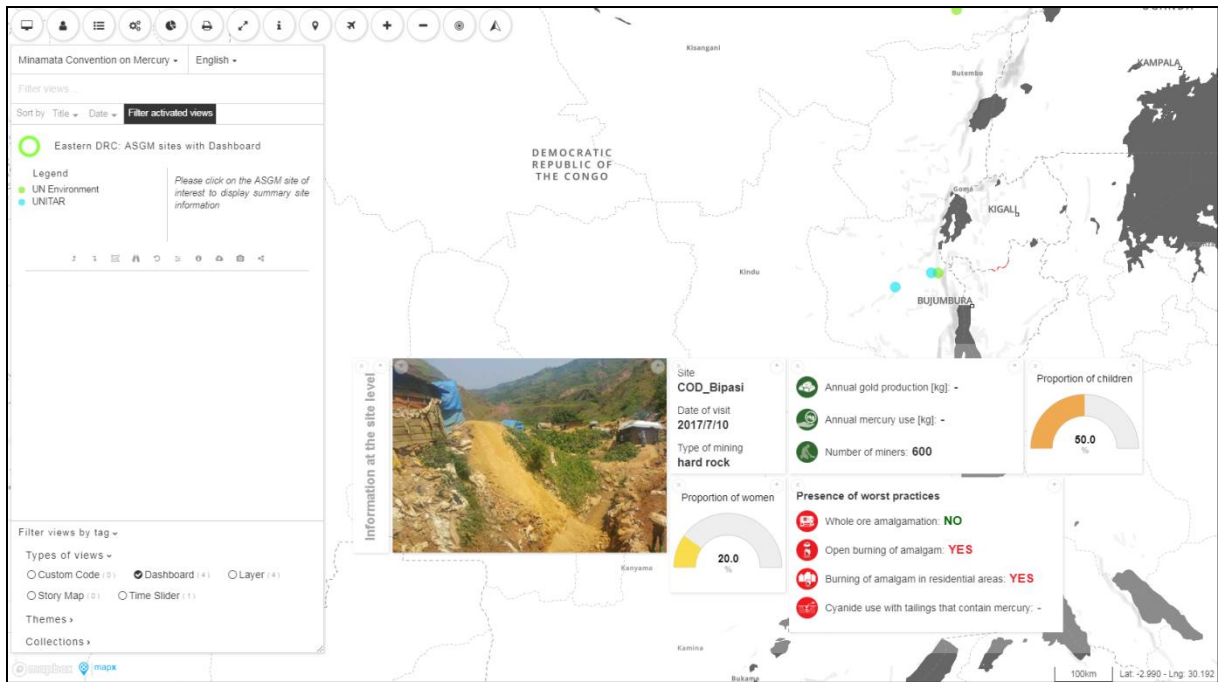


Figure 3: Exemple d'un dashboard sur la couche "Eastern DRC ASGM sites" dans le projet "Minamata Convention on Mercury".

Les story maps sont des présentations formées par une succession de slides, dans lesquels peuvent être insérés différents types de contenus comme du texte, des images ou des vidéos. Couplées à la visualisation de données sur la carte, ces informations permettent aux story maps d'avoir un impact visuel et dynamique. Les story maps sont utilisées pour présenter des projets de recherche, pour informer le public sur une problématique précise ou peuvent servir de support pour une présentation lors de conférences.

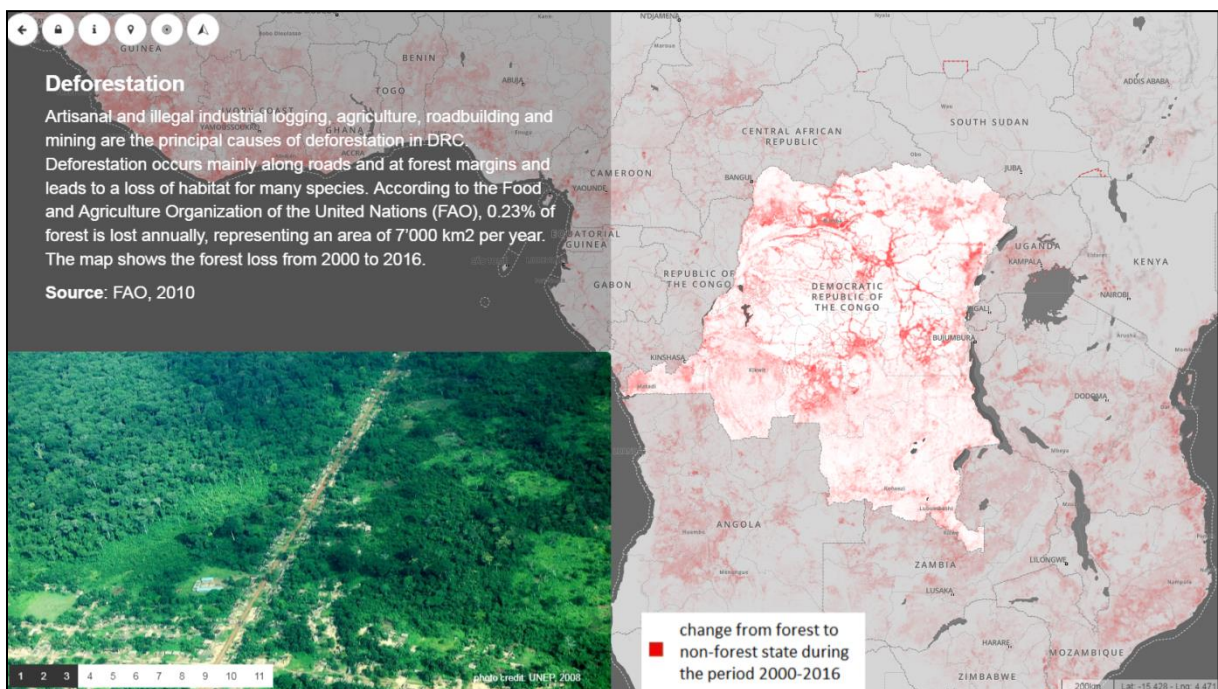


Figure 4: Exemple d'une slide extrait de la story map "Impact of mining activity on the forest cover" dans le projet "UN Environment Democratic Republic of Congo".

1.2.4. Source et intégrité des données

Les sources des données publiées dans MapX peuvent être diverses. Elles peuvent provenir directement des partenaires et clients de MapX (gouvernements, organisations non gouvernementales, instances internationales) ou alors d'autres institutions de recherche, de collecte et d'analyse de données dont les ressources sont accessibles librement.

Les données publiées dans MapX doivent être évaluées par le processus d'évaluation de l'intégrité des données (DIAF pour Data Assessment Integrity Framework). Toutes les couches qui sont ensuite publiées sous forme de vue doivent idéalement passer par ce processus pour s'assurer de leur validité, leur origine et leur pertinence. Les métadonnées doivent être remplies manuellement par l'utilisateur qui publie les couches dans MapX. Dans le cas d'une couche raster publiées par WMS, un lien vers les métadonnées doit être fourni.

Toutes ces étapes sont indispensables afin de s'assurer que les données publiées dans MapX soient exactes et permettent d'appuyer des prises de décision politique que cela soit dans un pays ou à l'échelle internationale.

1.3. Objectifs du stage

Ce stage a pour objectif de mettre en pratique les compétences acquises lors des cours du certificat complémentaire en géomatique dans différents projets concrets menés par le GRID – Genève et plus particulièrement par l'équipe MapX. Avec la multiplication des projets auxquels MapX apporte son soutien, il est nécessaire de pouvoir répondre aux besoins des institutions clientes, besoins souvent divers et qui peuvent évoluer avec le temps.

Ces besoins sont fluctuants en fonction du niveau de maîtrise des outils de SIG des personnes collaborant à ces projets, mais peuvent également varier en fonction de leur avancement (recherche de données, publication des couches sources, traitement et analyse des données, communication). Ce stage a donc permis de participer à plusieurs projets et à effectuer les tâches diverses énoncées ci-dessus.

Une première mission a consisté à collaborer avec l'équipe du Geneva Water Hub, pour créer un jeu de données spatiales concernant les conflits armés impactant les ressources en eau à travers le monde. Ce jeu de données a servi ensuite de support pour une story map présentée lors de la Geneva Peace Week 2018.

Deux missions ont été consacrées à l'écriture de scripts R permettant l'automatisation de différents processus de traitement de données en vue de leur publication dans MapX. Le premier script permet aux formulaires créés via l'application KoBoCollect d'être publiés dans MapX. KoBo est une application mobile qui permet de récolter des données humanitaires sur le terrain via des formulaires personnalisables. Le deuxième script a pour but la publication et la mise à jour dans MapX des données de Famine Early Warning Systems Network (FEWS NET). Ces données, qui sont actualisées 3 fois par an, font état du niveau d'alerte alimentaire de certaines régions vulnérables (Asie, Afrique, Amérique latine). La mise à jour de ces données permet donc de présenter l'état le plus récent de la situation humanitaire sur le terrain.

Une autre mission a été la création d'un document de synthèse relatif au secteur minier artisanal dans la partie est de la République Démocratique du Congo (RDC) d'après le rapport 2015 publié par

l'International Peace Information Service (IPIS). Ce rapport, basé sur des campagnes de terrain, recense et localise les sites miniers artisanaux et fournit bon nombre d'informations sur chaque site, comme le nombre de travailleurs, le type de minéraux exploités, la présence de groupes armés etc. Ce document a par la suite été inclus comme la partie introductive d'un article scientifique portant sur la détection des activités minières illégales par imagerie satellite nocturne, grâce aux émissions lumineuses.

Une autre mission ponctuelle a été le test d'une nouvelle fonctionnalité de MapX qui permet le calcul de l'aire d'intersection entre plusieurs couches affichées à l'écran. La méthode de calcul est différente de ce qui se fait avec un logiciel de SIG, le but est donc ici de comparer les deux méthodes et de voir si le calcul effectué dans MapX est pertinent.

Finalement, une partie du temps de ce stage a été consacré à la publication de données et à la création de vues dans MapX pour certains projets, mais également à la traduction de tutoriels pour l'utilisation de MapX de l'anglais vers le français. Ces tutoriels sont employés par l'équipe MapX lors de sessions de "training" auprès des personnes appelées à utiliser la plateforme.

2. PROJETS

2.1. Collaboration avec le Geneva Water Hub

2.1.1. Présentation

Le [Geneva Water Hub](#) est une initiative conjointe de la Direction du développement et de la coopération de la Confédération suisse et de l'Université de Genève visant à favoriser la mise en place d'un agenda en matière d'hydro-politique au niveau global et l'aide à la coopération entre les différents acteurs en ce qui concerne les enjeux liés à l'eau.

Pour ce faire, le Geneva Water Hub axe sa démarche à travers trois fonctions principales qui sont premièrement la recherche et l'éducation (conduite d'analyses stratégiques afin de permettre la prise de décision basée sur des preuves scientifiques), deuxièmement un pôle de compétences en hydro-politique pour promouvoir la mise en place de politiques et finalement le réseautage dans le but de mettre en relation les organisations et acteurs divers actifs sur le thème de l'eau.

Dans le cadre de la Geneva Peace Week 2018, le Geneva Water Hub participe à une journée de conférences sur le thème "Water Diplomacy: Building Bridges for Lasting Peace". La collaboration avec MapX a été mise en place en vue de proposer une présentation appuyée par des données géospatiales directement visualisées dans la plateforme via une story map.

L'objectif de cette collaboration a été de montrer les différents impacts sur l'eau que peuvent avoir les conflits armés, notamment au niveau de l'accès à l'eau potable pour les populations civiles ainsi que le contrôle d'infrastructures stratégiques et leur endommagement.

2.1.2. Méthode

La première étape de ce travail a été la recherche des dépêches de presse relatant des événements de conflits ayant eu un impact direct ou indirect sur l'eau. Le site [OOSKANews](#), site d'information spécialisé dans les problématiques liées à l'eau a été choisi comme base à partir de laquelle une première sélection a été opérée au niveau temporel, soit la période 2012 – 2018, avec la grande majorité des événements s'étant produits entre 2015 et 2018. Plusieurs pays ont été sélectionnés afin de montrer la diversité et la complexité des conflits et leur impact sur l'accès à l'eau: la Colombie, l'Irak, la Syrie, le Yémen, la Lybie, Gaza et l'Ukraine.

Une table de données a ensuite été mise sur pied pour classer ces événements selon des critères objectifs et applicables à toute situation. Ces informations peuvent être regroupées selon plusieurs types.

- Géographiques: Le pays, la division administrative immédiatement inférieure (district, région, département, gouvernorat), la ville la plus proche, ainsi que les coordonnées de latitude et de longitude.
- Qualitatives: Le type de problème (conflit armé, infrastructures touchées, impact humanitaire, problème d'ordre général, problème de santé comme une épidémie), le type d'évènement (description plus détaillée du problème, comprenant par exemple le type de conflit armé, si l'impact humanitaire est positif ou négatif, si l'accès à l'eau a été coupé, est revenu à la normal, si des pollutions ont été engendrées). D'autres informations sont également présentes comme le type d'infrastructure touchée (barrage, station d'épuration,

pipeline) ainsi que les éventuels dégâts suite à l'évènement (site contaminé, usine détruite etc..).

- Des informations quantitatives sont également fournies quant au nombre de personnes touchées par l'attaque.
- La source de l'article à partir duquel les informations ont été extraites ainsi qu'une description détaillée de l'évènement peuvent être remplies si nécessaire.

Finalement, les données ont été publiées sur MapX et une story map a été créée. Plusieurs couches de données ont été nécessaires afin de placer les données concernant les conflits dans un contexte clair. Ces couches sont les suivantes:

- Vues satellites ([MapBox](#), [Sentinel-2](#));
- Couche globale des rivières principales ([Main global drainage system \(TWAP\)](#));
- Densité de population ([Gridded Population of the World v4](#));
- Couche globale des barrages et réservoirs ([Global Reservoir and Dam \(GRanD\) v1](#));
- Cas de choléra par gouvernorat au Yemen pour le mois d'octobre 2017 ([Yemen: Cholera Outbreak Epidemiology Update Data](#));
- Frontières internationales ([Country Borders](#));

2.1.3. Résultats

Le jeu de données initiales comporte 84 évènements ayant eu un impact sur l'accès à l'eau dans un contexte de conflit armé. La classification du type de conflit différencie les conflits armés, les problèmes d'ordre général (déclarations de politiciens, d'instance gouvernementales et internationales), les problèmes concernant la santé (propagation de maladies, épidémies), les urgences humanitaires et les évènements impactant directement les infrastructures.

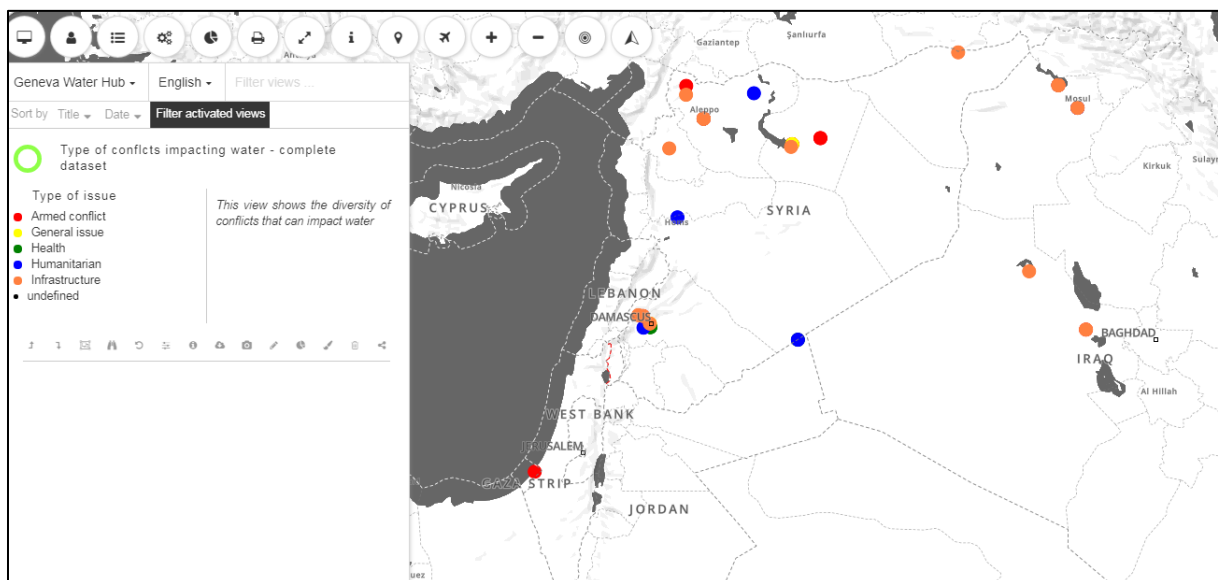


Figure 5: Vue MapX de la couche des évènements de conflits armés, zoom sur l'Irak, la Syrie et la Palestine.

Parmi ces 84 évènements, une sélection a été opérée pour inclure dans la story map neuf cas représentant la variété des impacts sur l'eau que peuvent avoir les conflits armés. Les informations contenues dans le jeu de données spécifique à la story map sont identiques à la différence de la classification du type de conflits, ici définie selon les critères de la [Geneva Academy RULAC](#), projet

qui vise à mettre en lumière les problèmes liés aux conflits armés en regard du droit humanitaire international.



Figure 6: Liste des évènements présentés dans la story map.

Une fois les données publiées dans MapX, la story map a pu être élaborée en concertation avec toutes les personnes impliquées. La story map devant servir de support à la présentation lors de la Geneva Peace Week, une narration présentant des cas concrets de problèmes liés à l'eau en parcourant différents pays a été choisie. Pour chaque pays, un zoom national présente la situation générale, les forces en présence et le type de conflit en cours ou passé.

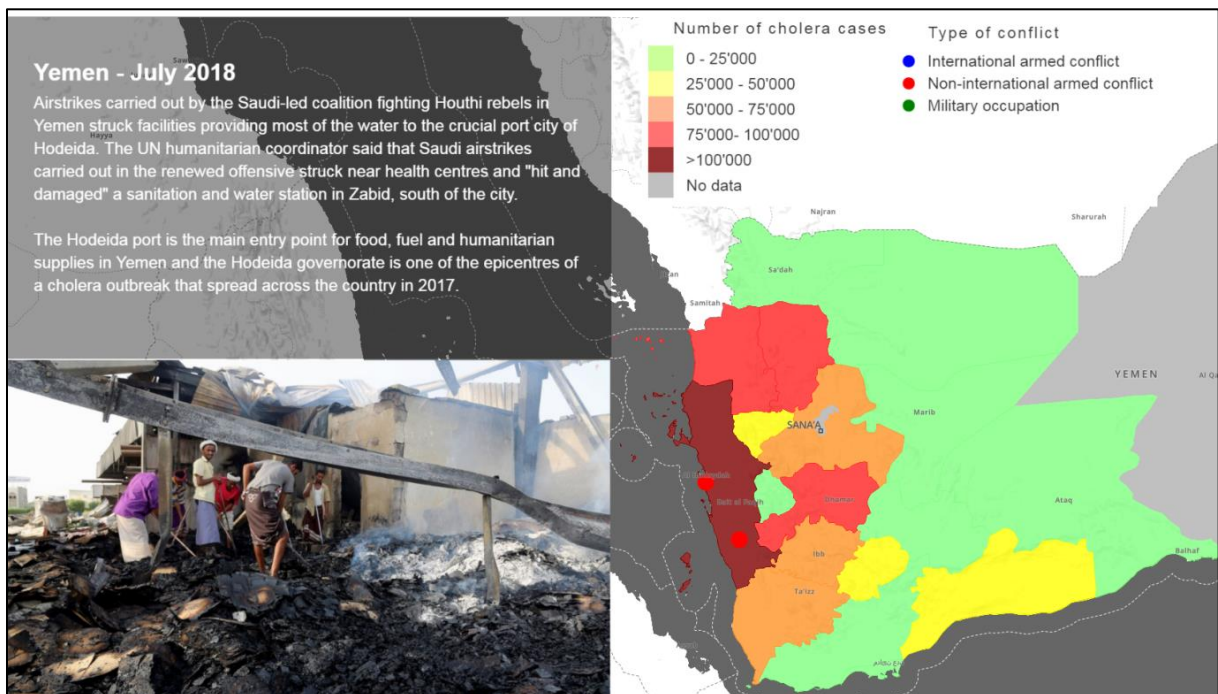


Figure 7: Exemple issu de la story map concernant un évènement au Yemen en juillet 2018.

2.1.4. Bilan

Ce projet a été l'occasion d'appréhender les possibilités mais également les limites relatives à l'utilisation de données géospatiales dans le cadre d'une collaboration avec d'autres institutions. Le

but d'une telle collaboration avec le Geneva Water Hub était de fournir un support visuel, interactif, basé sur des données fiables pour une présentation traitant du droit international humanitaire, plus particulièrement sur la thématique de l'eau.

Cette collaboration avec le Geneva Water Hub a permis d'étendre la visibilité de MapX, notamment grâce à la story map utilisée comme support pour la présentation de Mara Tignino lors de la conférence de haut niveau durant la Geneva Peace Week. Le fait de conduire ce projet depuis les premières étapes jusqu'à la concrétisation et la communication a également été positif.

Tout au long de ce projet, le défi majeur a consisté à collaborer avec des acteurs ne possédant que peu ou pas du tout de connaissances techniques dans l'utilisation des SIG et ayant parfois du mal à saisir la nature même de leurs possibilités. Il a en effet été parfois difficile de trouver un terrain d'entente sur la marche à suivre et sur la finalité du projet en lui-même.

2.2. Intégration des données KoBoCollect dans MapX

2.2.1. Présentation

[KoBoCollect](#) (KoBo) est une application mobile open source créée et développée par Harvard Humanitarian Initiative (HHI), Brigham and Women's Hospital (BWH) et KWEYO, avec le soutien d'institutions rattachées aux Nations Unies telles que le Bureau de la coordination des affaires humanitaires (OCHA), l'Agence des Nations Unies pour les réfugiés (UNHCR). Cette application permet la collecte de données sur le terrain via des formulaires personnalisables.

Les réponses au formulaire peuvent être enregistrées hors ligne, et être compilées dans un seul fichier lorsqu'une connexion internet est disponible. Un atout majeur de cette application est la possibilité qu'un nombre illimité de personnes puissent participer à la collecte de la donnée (en utilisant le même formulaire), et qu'un seul fichier regroupant toutes les soumissions soit créé en sortie.

Cette application est très utilisée dans le domaine humanitaire car elle permet une collecte rapide et fiable de données, basées sur des observations faites directement sur le terrain. Des exemples très concrets d'utilisation de KoBo à citer sont par exemple le recensement des populations installées dans des camps de réfugiés, tant au niveau quantitatif (nombre d'habitations, nombre de personnes occupant ces habitations), qu'au niveau qualitatif (accès à l'eau potable, à la nourriture, conditions de sécurité etc...) ou alors l'évaluation de l'état de maintenance du réseau électrique de localités en Haïti, afin de planifier la reconstruction et l'amélioration de ce réseau.

Après la récolte des données sur le terrain, il est possible d'exporter un fichier compilant toutes les entrées sous forme d'un tableau (Excel, CSV) dans lequel chaque colonne représente une question et ou chaque ligne correspond à un formulaire rempli. De cette manière, toutes les informations sont centralisées et sont exploitables par la suite.

Si les données récoltées contiennent des informations spatiales (latitude et longitude), il peut être intéressant de les visualiser sur une carte. Pour réaliser une telle opération, l'utilisateur doit posséder des connaissances en SIG, ce qui n'est pas toujours le cas et peut représenter un problème. Ce script permet donc de pallier ce manque, en offrant la possibilité à l'utilisateur de publier les données dans MapX beaucoup plus facilement.

2.2.2. Méthode

R a été choisi comme langage de programmation pour ce script car il possède les avantages d'être compatible avec MapX et d'offrir une grande souplesse pour le traitement de données. Ce script utilise une requête cURL (client URL request library) pour se connecter à la base de données de l'API KoBo. Cette requête demande à l'utilisateur d'entrer son identifiant KoBo ainsi que son mot de passe, et le numéro (id) du formulaire désiré. Si ce numéro n'est pas connu de l'utilisateur, une fonctionnalité permet d'afficher une liste contenant les informations sur tous les formulaires de l'utilisateur.

Un dataframe est ensuite créé dans R pour lire les données. A partir de ce dataframe, plusieurs opérations sont effectuées dans le but de rendre les données plus propres en vue de leur incorporation dans MapX. Ces opérations incluent notamment la conversion des champs de latitude et longitude en nombres, afin d'être utilisables dans une carte. D'autres opérations sont également réalisées, comme la recherche d'éventuelles entrées ne contenant pas d'indication de longitude ou de latitude, entrées qui sont effacées car ne pouvant pas être représentées sur une carte.

Une fois les données stockées dans la base de données PostgreSQL de MapX sous forme de couche source, l'utilisateur doit manuellement remplir les métadonnées et renseigner sur l'intégrité de ces données. Cette étape ne peut pas être automatisée, mais elle ne doit être effectuée qu'une seule fois, lors de la publication initiale. Le script permet en effet d'effectuer le processus selon une fréquence désirée (quotidienne, hebdomadaire, mensuelle etc...), ce qui peut se révéler utile si un utilisateur désire visualiser l'évolution de la collecte de données avec le temps.

2.2.3. Résultats

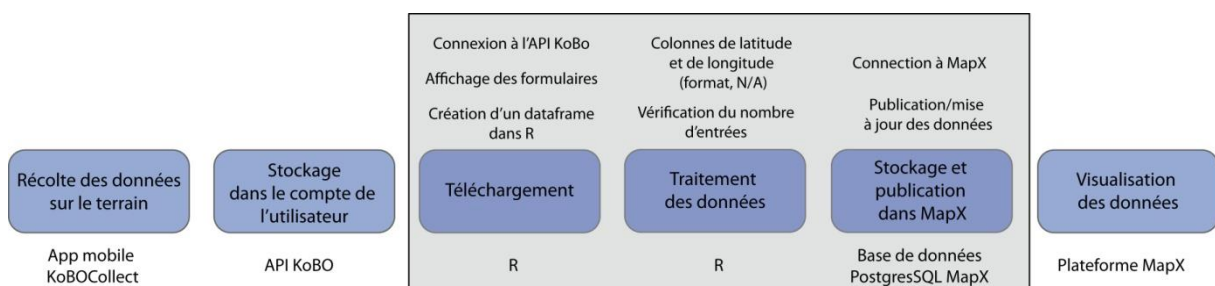


Figure 8: Processus de traitement des données de KoBo vers MapX.

La Figure 8 ci-dessus montre un schéma explicatif du processus complet partant de la récolte des données sur le terrain via l'application mobile KoBo jusqu'à la visualisation des données dans la plateforme MapX. Le rectangle gris met en évidence les tâches qui ont pu être entièrement ou partiellement automatisées, et qui peuvent être répétées grâce à une exécution périodique du script.

2.2.4. Bilan

Ce projet était initialement prévu pour être implémenté dans l'API KoBo directement et permettre d'exporter les données depuis l'interface de l'utilisateur vers MapX (en les convertissant au format GeoJSON par exemple). Face au manque de communication et de réponse de la part de l'équipe KoBo, l'approche inverse a été choisie, c'est-à-dire que le script a été conçu pour aller chercher les données dans KoBo à partir de MapX.

Un utilisateur qui désire visualiser ses données devra donc d'abord posséder un projet dédié dans MapX, remplir les métadonnées et créer une vue à partir des données sources, étape qui est de toute manière indispensable pour chaque vue dans MapX, peu importe l'utilisateur, le projet ou les données. L'automatisation du processus permet en revanche de gagner beaucoup de temps par la suite, si les données de base viennent à être mises à jour.

Ce script est un outil qui simplifie la visualisation de données à la fois durant la campagne de terrain mais aussi ultérieurement lors de la phase d'analyse. La possibilité de visualiser les données déjà récoltées permet à l'utilisateur d'avoir un aperçu de l'avancement de son projet et ainsi de planifier efficacement la suite de son travail sur le terrain.

Une fois la récolte terminée, l'utilisateur a la possibilité de visualiser ses données dans un contexte plus large, en employant tous les outils disponibles dans MapX. Ce script est un moyen pour les utilisateurs sans connaissance préalable des logiciels SIG de travailler avec des données géospatiales, il peut être employé au cas par cas.

Une limite de ce script concerne la géométrie des données à importer dans la base de données de MapX, en effet il n'est actuellement pas possible de travailler sur des données de lignes ou de polygones. Cette limite n'est pas un problème en soit puisque les données récoltées avec KoBo sont presque exclusivement des points, elle peut néanmoins se révéler problématique dans certaines situations spécifiques.

2.3. Intégration des données FEWS NET dans MapX

2.3.1. Présentation

[Famine Early Warning System Network](#) (FEWS NET) est un système permettant de rendre compte du niveau d'insécurité alimentaire dans certaines régions du monde. Ce réseau a été constitué par l'USAID en 1985 dans le but d'aider à la planification de crises alimentaires en se basant sur des données factuelles. Les régions recensées sont au nombre de 5 et incluent l'Asie centrale, l'Amérique centrale et les Caraïbes, l'Afrique de l'est, l'Afrique de l'ouest et le sud de l'Afrique.

Ces données sont actualisées 3 fois par an (février, juin et octobre) avec des ajouts ponctuels de prévisions à court et moyen terme concernant l'évolution de la situation pour une région donnée. La régulière mise à jour de ces données permet donc une planification efficace pour les programmes humanitaires.

Les données sont accessibles librement depuis le FEWS NET [Data Center](#) et peuvent être téléchargées sous forme d'images ou sous forme d'ESRI Shapefile. Le niveau d'insécurité alimentaire est représenté par des valeurs allant de 1 (niveau d'alerte minimal) à 5 (famine). Les endroits dans lesquels le niveau d'alerte serait plus élevé sans aide humanitaire sont mis en évidence par des points d'exclamation (Figure 9).

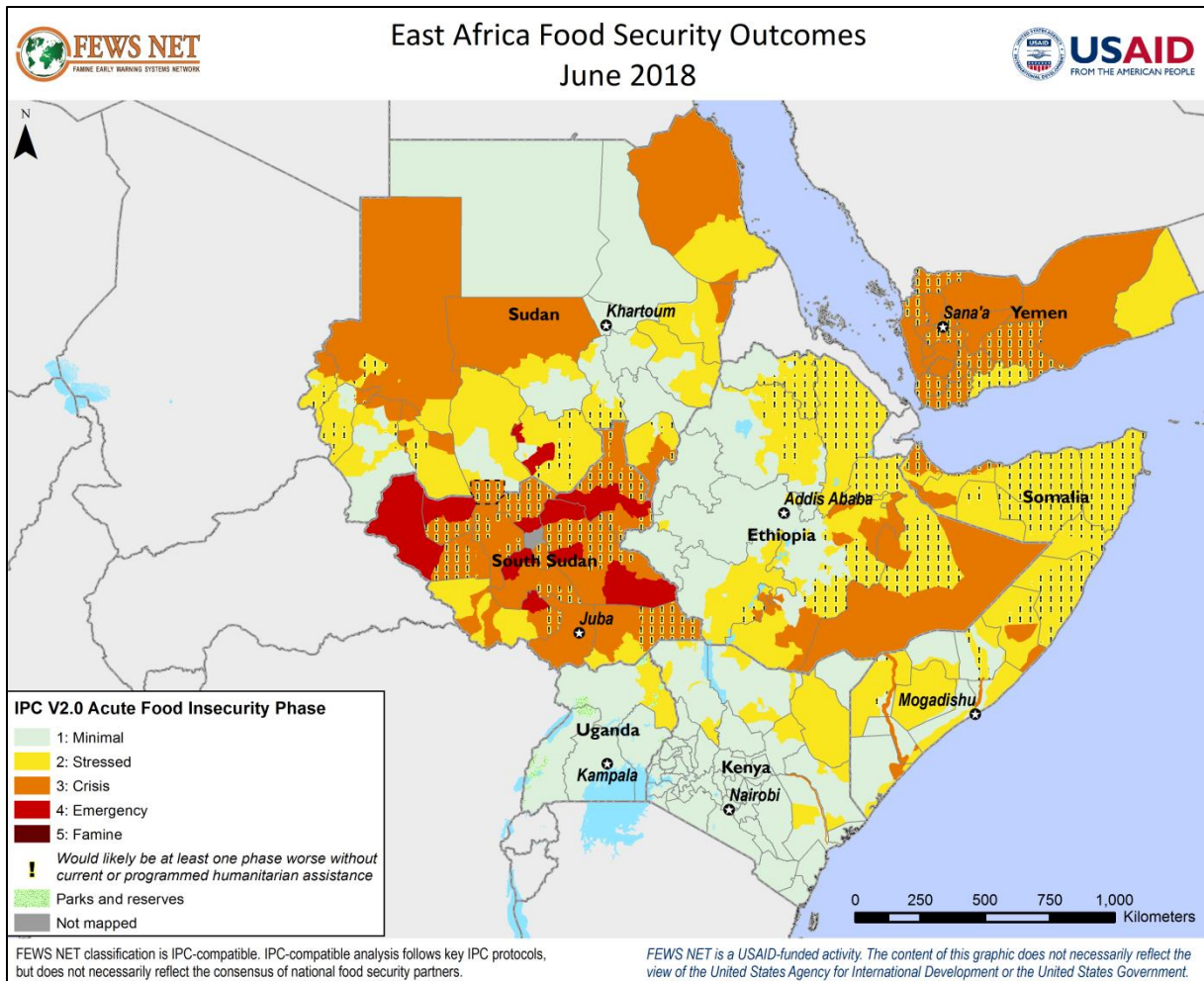


Figure 9: Exemple de la situation pour la région d'Afrique de l'est en juin 2018.

Actuellement ces données sont publiées dans MapX pour certains pays. Leur régulière mise à jour demande néanmoins que les couches soient republiées à chaque fois, soit plusieurs fois par an. Cette tâche peut s'avérer chronophage et répétitive, d'où le besoin d'un script permettant l'automatisation du téléchargement, du traitement et de la publication des données dans MapX.

2.3.2. Méthode

Le script a été programmé en R car ce logiciel possède des packages performants dans le traitement de données spatiales (sf et geosonio ont été utilisés dans le cas présent), en plus des raisons énoncées au point 2.2.2. Le but du script est de télécharger les données sous forme de dossier compressé contenant les Shapefile de chaque région, d'ouvrir ces Shapefile dans R, de les traiter puis de les grouper dans une couche globale puis de les stocker dans la base de données PostgreSQL de MapX. Une option supplémentaire pour créer et exporter cette couche globale sous forme de GeoJSON a également été mise en place.

L'étape de traitement des données est ici importante et concerne la conversion des couches de "MultyPolygon" en collection de polygones séparés. La structure des données téléchargées depuis FEWS NET Data Center possède en effet une géométrie très simple qui regroupe tous les polygones par niveau d'alerte, c'est-à-dire qu'en cliquant sur un polygone ayant une valeur d'alerte minimum (valeur de 1), tous les polygones situés dans la même région et de la même valeur seront

sélectionnés, même s'ils ne se touchent pas. Cette structure n'est pas souhaitable si l'on veut effectuer de l'analyse spatiale sur des polygones séparés, ou sur une extension spatiale donnée.

2.3.3. Résultats

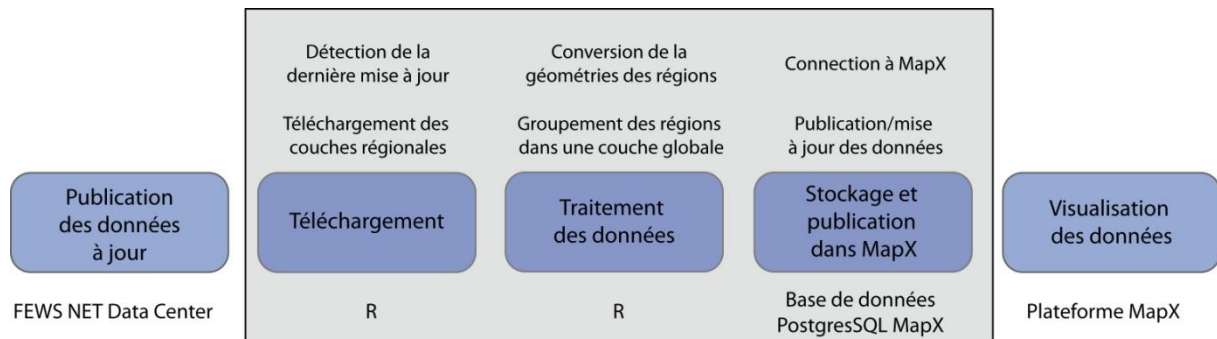


Figure 10: : Processus de traitement des données de FEWS NET vers MapX.

La Figure 10 ci-dessus montre un schéma explicatif du processus complet partant de la mise à jour des données sur le FEWS NET Data Center jusqu'à la visualisation des données dans la plateforme MapX. Le rectangle gris met en évidence les tâches qui ont pu être entièrement ou partiellement automatisées et qui peuvent être répétées grâce à une exécution périodique du script.

2.3.4. Bilan

Ce script permet de faciliter la publication des données à jour d'alerte de famine dans MapX, ce qui rend ces données plus pertinentes et facilement accessible et visualisable dans la plateforme. Le regroupement des différentes régions en une seule permet ainsi de facilement d'avoir un aperçu global et non plus fragmenté de la situation la plus récente.

La conversion du type de géométrie permet également de travailler aisément sur des polygones séparés plutôt que sur une liste de plusieurs polygones qui partagent la même valeur de niveau d'alerte. Les données sont ainsi utilisables de manière plus souple qu'avec la structure initiale.

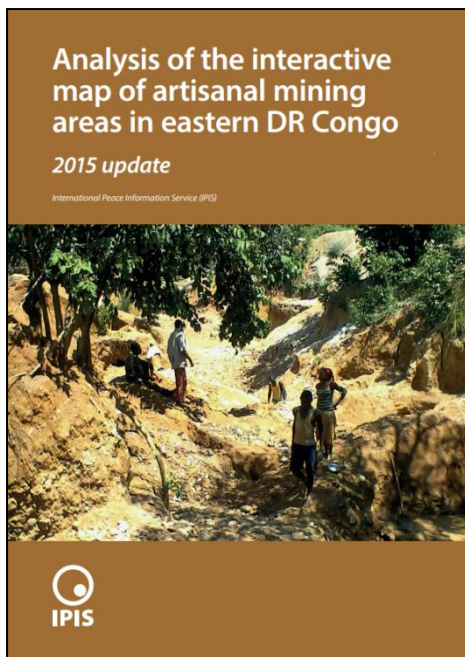
Le gain de temps réalisé est également conséquent. Grâce à ce script, il n'est en effet plus nécessaire de s'occuper manuellement de la publication et de la mise à jour des données, ni de leur agrégation en une couche unique. La publication de la couche source dans MapX et son actualisation pouvant en effet prendre un certain temps et se révéler redondante.

2.4. Travail de recherche sur le secteur minier artisanal en RDC

2.4.1. Présentation

[IPIS](#) est un institut basé à Anvers en Belgique qui œuvre dans la production et la recherche de données permettant de soutenir les acteurs qui travaillent à favoriser la paix, le développement durable et le respect des droits de l'homme. IPIS travaille particulièrement en Afrique subsaharienne.

Une des missions phares menées par IPIS concerne la cartographie de l'exploitation minière artisanale dans l'est de la RDC, secteur précaire et non régulé par l'état, pourtant source de revenu pour des millions de personnes. Dans ce cadre, d'importantes campagnes de terrain sont organisées



régulièrement afin de recenser les sites miniers artisanaux et les informations qui leur sont relatives comme le nombre de travailleurs, les minéraux exploités, l'éventuelle présence de groupes armés qui contrôlent l'activité, l'utilisation du mercure pour récupérer l'or, et le contrôle des sites par l'état.

Toutes les informations sont ensuite publiées dans une [carte interactive](#), accompagnée d'un [rapport](#) explicatif, le dernier ayant été publié en 2016 et concerne les données récoltées jusqu'en 2015. Ce rapport présente dans leur globalité les résultats des campagnes de terrain et contextualise les données dans le cadre plus large des problèmes de gouvernance, de corruption, de violence et de problèmes environnementaux.

Figure 11: Rapport IPIS 2015 (IPIS 2016)

2.4.2. Méthode

Dans le but d'apporter un contexte plus général aux données IPIS, un travail de recherche bibliographique a été réalisé sur le thème du secteur minier artisanal dans l'est de la RDC et plus particulièrement sur les problèmes que soulève cette activité. Ont pu être identifiées quatre problématiques principales, qui sont le manque de contrôle de la part de l'état, la présence de groupes armés, les problèmes environnementaux et finalement les problèmes sociaux liés à l'activité.

Cette recherche bibliographique s'accompagne d'une synthèse concernant les informations récoltées par les différentes campagnes de terrain. Les sites miniers artisanaux sont représentés sur la carte interactive en fonction du minéral principalement exploité, mais d'autres informations sont aussi répertoriées et peuvent être utilisées comme critères de sélection. Une liste contenant une courte description de chaque attribut a donc été dressée.

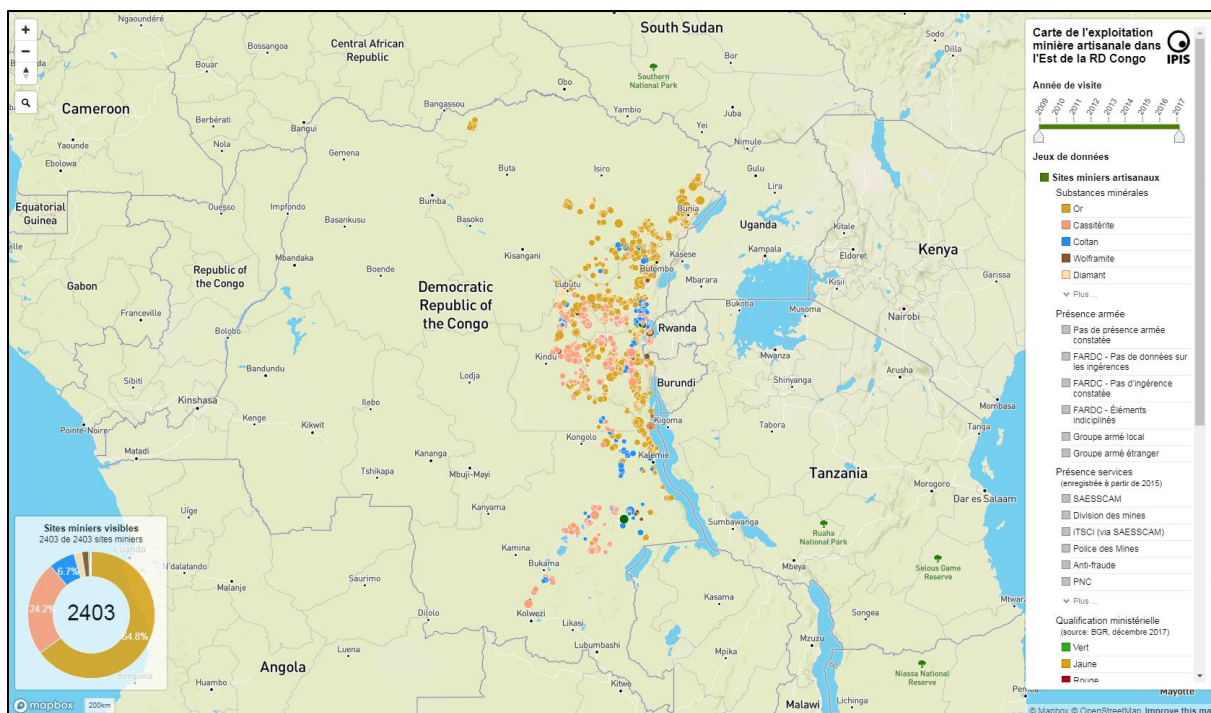


Figure 12: Capture d'écran de l'interface présentant la carte interactive créée par IPIS. Sur la droite, un menu permet de filtrer les sites miniers artisanaux selon des critères prédéfinis.

2.4.3. Résultats

L'activité minière artisanale est un secteur permettant à des millions de personnes de vivre, tout en demeurant source de problèmes sociaux, environnementaux et sécuritaires. Quatre principales problématiques liées au domaine minier artisanal ont pu être mises en évidence.

- Le manque de contrôle de l'état. L'instabilité politique de la RDC suite notamment aux guerres civiles de 1996-1997 et 1998-2003 a permis à la corruption de se développer de manière importante dans le pays, mais a également favorisé l'émergence de secteurs d'activités illégales, dont fait partie l'exploitation minière artisanale.
- La présence de groupes armés. Profitant du manque de contrôle étatique, de nombreux groupes armés se sont développés dans la région de l'est de la RDC, pratiquant le racket et l'extorsion envers les mineurs. Cette activité très lucrative est en vigueur dans la majorité des sites artisanaux et s'exerce à des degrés divers par des groupes armés locaux, étrangers ou même par les forces armées corrompues de l'état.
- Les problèmes environnementaux. Mise à part la dégradation du paysage par une activité minière non régulée, l'enjeu environnemental principal dans la région concerne l'utilisation de mercure pour la récupération de l'or par processus d'amalgamation. Le mercure est en effet un métal hautement toxique pour les organismes qui se trouve en concentrations élevées lorsqu'il est rejeté dans l'environnement.
- Les problèmes sociaux. Lorsque l'on parle de la RDC, on évoque souvent la "malédiction des ressources naturelles". Ce concept explique le lien entre la présence d'abondantes ressources naturelles et les conflits qu'elles engendrent. Les mineurs travaillant dans des mines artisanales sont souvent employés dans des conditions difficiles, victimes des groupes armés, et évoluant dans des conditions de grande insécurité.

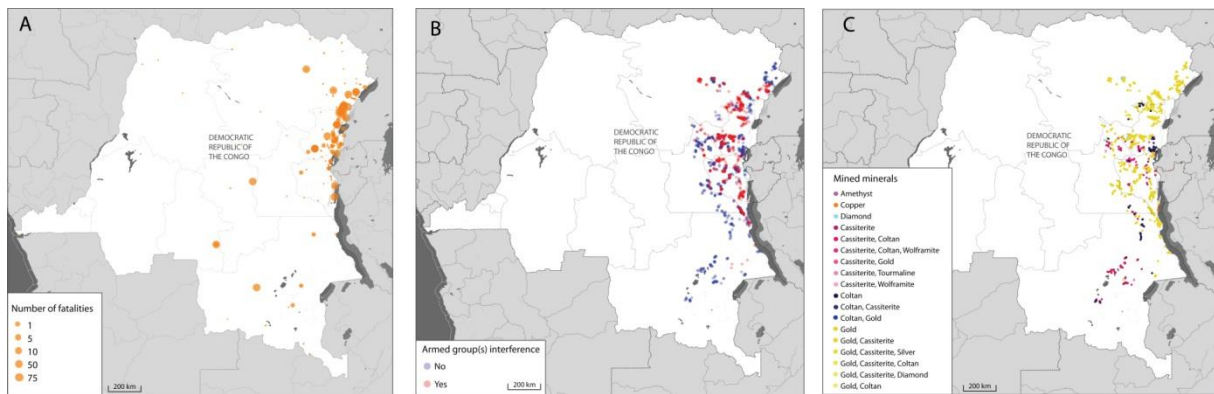


Figure 13: Vues MapX montrant A) la localisation des conflits (ACLED), B) L'interférence de groupes armés dans les sites miniers artisanaux (IPIS) et C) Les sites miniers artisanaux classés par minéral principal (IPIS).

Les données présentées sur la carte interactive d'IPIS permettent de se rendre compte de toutes ces problématiques car en plus des différents minéraux exploités, du nombre de travailleurs présents, et d'autres informations générales sur la mine, sont indiqués des renseignements concernant la présence/absence d'un ou plusieurs groupes armés, l'utilisation éventuelle de mercure ainsi que la présence/absence d'organes de surveillance.

2.4.4. Bilan

Etudier ces données dans un cadre plus large a permis notamment de se rendre compte de leur pertinence dans le sens qu'elles permettent d'apporter des informations importantes et factuelles sur des situations complexes. On voit ici un très bon exemple de données pouvant servir de support pour la mise en œuvre d'une politique nationale permettant d'intégrer l'activité minière artisanale dans un cadre plus régulé.

D'autre part, cette mission a été l'occasion de mener une recherche bibliographique sur un sujet complexe, au sein duquel beaucoup de variables entrent en ligne de compte. Cette recherche a permis également de mettre en évidence l'impact que peuvent avoir des données géospatiales dans un contexte plus large.

2.5. Tests des calculs de l'aire d'intersection avec l'outil PixOp

2.5.1. Présentation

Une nouvelle fonctionnalité de MapX actuellement en cours de développement est un outil qui permet de calculer la surface d'intersection des couches en cours de visualisation en prenant comme étendue de calcul l'étendue de ce qui est affiché à l'écran grâce à PixOp. Cet outil fonctionne sur les couches vectorielles et effectue le calcul d'intersection en fonction d'une grille dont la résolution dépend du niveau de zoom, et non pas sur les géométries précises inhérentes aux couches, comme cela se fait dans un SIG standard.

Pour cette raison, le calcul effectué dans MapX est susceptible de différer de celui effectué avec un SIG comme QGIS. Par ces tests, l'objectif était de voir si la différence entre ces deux calculs est suffisamment faible pour utiliser l'outil de MapX sans restriction, ou alors si certaines limites entravent son utilisation, et si oui, lesquelles.

2.5.2. Méthode

Plusieurs tests ont été effectués sur des overlaps de couches à différentes échelles, avec des géométries plus ou moins complexes et avec différents niveaux de zoom, en comparant les résultats obtenus dans MapX et dans QGIS. Les différents calculs d'aires des intersections ont été faits avec des couches en WGS84 dans QGIS et dans MapX. 6 situations d'overlaps ont été testées dans 4 pays avec des superficies variables:

- Haïti (superficie = 27 750 km²) → 1 situation;
- Somalie (superficie = 637 657 km²) → 2 situations;
- République Démocratique du Congo (superficie = 2 345 409 km²) → 1 situation;
- Brésil (superficie = 8 514 876 km²) → 2 situations

Pour chaque situation, le calcul d'aires d'intersection a été opéré sur deux couches en faisant varier le niveau de zoom (en partant du niveau le plus proche défini par l'outil "zoom to all features"), mais également en faisant varier virtuellement la résolution (scale) de l'écran. Des couches avec des niveaux de complexité différents ont été employés afin de voir dans quelle mesure cette complexité impacte sur la précision du calcul. Les couches utilisées sont les suivantes:

- Dans le projet "World", les couches "Somalia flood – May 2018 (OCHA Somalia – 2018)", "Biodiversity hotspots (Conservation International, 2016)", et "Post-flooded areas over Jowhar Middle Shabelle region (UNOSAT – SWALM, 2015)".
- Dans le projet "UN Environment Democratic Republic of Congo", la couche "DRC Mining Cadastre (Ministry of mining, 2016)".
- Dans le projet "UN Environment Haïti", la couche d'inondations (Centre National de l'Information Géo-Spatiale)"
- Dans le projet "UN Biodiversity Lab – Brazil", la couche "Intact Forest Landscapes (IFLs)", "Special programme projects in Latin America and the Caribbean (UN ENvironment, 2018)" et "Marine Ecoregions of the World (MEOW)".
- Deux couches ont été manuellement uploadées sur MapX, la première est une couche issue de FEWS NET (avec uniquement le niveau d'insécurité "stressed" pour juin 2018) clippée pour Haïti, la deuxième est la couche des aires protégées "IUCN Green List of Protected and Conserved Areas" préalablement traitée dans QGIS pour éviter des problèmes de superposition (Dissolve).

2.5.3. Résultats

Le document détaillant les tests effectués est donné en annexe (ANNEXE X) et présente les résultats pour chaque situation. Sont brièvement présentés ici les principaux résultats obtenus. La Figure 14 montre les différents niveaux de complexité des couches testées allant d'une géométrie très simple (Fig. 14A Somalie), moyenne (Fig. 14B Haïti) ou extrêmement complexe (Fig. 14C Somalie).

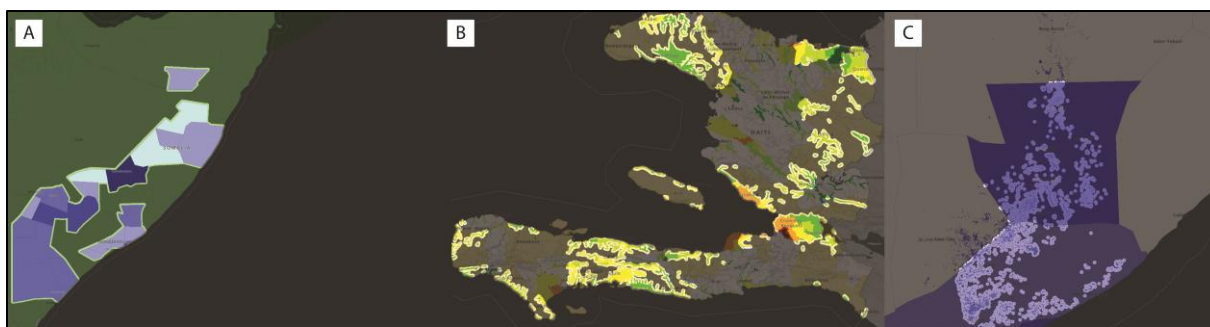


Figure 14: Comparaison du niveau de complexité des géométries de plusieurs couches.

La variable ayant le plus d'impact sur la précision du résultat est la complexité de la géométrie. Lorsque les géométries des couches sont très simples, la différence entre un calcul effectué avec QGIS et celui effectué dans MapX varie aux alentours de 1% et moins. Lorsque les géométries augmentent en complexité, cette erreur peut atteindre 12%, tandis que pour des cas extrêmes comme la Figure 14C, la seule différence entre les deux méthodes de calcul est elle-même supérieure au résultat obtenu dans MapX.

Le niveau de zoom est également une variable importante. Dans une première mesure, il influence directement la surface calculée (à cause de la simplification des géométries lorsque l'utilisateur dézoome) donc les polygones se verront approximés et réduits (l'aire calculée diminue toujours). Dans une autre mesure, le niveau de zoom peut être problématique si la géométrie de la couche est très complexe mais possède une large étendue.

La résolution de l'écran ("scale") influence également le résultat d'une manière limitée dans le sens qu'elle permet de se rapprocher de la valeur calculée dans un SIG, uniquement pour les géométries complexes. Pour la plupart des calculs effectués, passer d'une scale de 1 à 2 ne change en effet pas significativement le résultat.

2.5.4. Bilan

L'importance de tester une nouvelle fonctionnalité en faisant varier certains paramètres, afin de s'assurer de sa fiabilité a pu être mise en évidence grâce à cette mission. Ces tests ont permis de montrer les possibilités offertes par cet outil mais également ses limites.

La comparaison entre le résultat obtenu via un calcul dans QGIS et celui obtenu dans MapX montre que l'outil de MapX ne donne pas les garanties nécessaires pour une utilisation dans un but quantitatif. La variation de la différence entre les deux méthodes pouvant aller de moins de 1% jusqu'à plusieurs dizaines montre en effet que la complexité de la géométrie est un facteur qui influence trop les incertitudes pour que cet outil soit employé génériquement.

Cet outil pourrait néanmoins être utilisé dans un but de visualisation, pour mettre en évidence les zones de chevauchement entre plusieurs couches vectorielles, sans besoin de calculer l'aire d'intersection. L'utilisation d'un logiciel de SIG reste par conséquent nécessaire pour effectuer un calcul précis, surtout si les chiffres désirés doivent être utilisés par la suite.

2.6. Publication de données dans MapX

2.6.1. Présentation

Plusieurs étapes sont nécessaires avant d'arriver à la visualisation finale des données dans MapX. Chaque vue présente dans la plateforme est créée à partir d'une couche source qui peut être une couche raster ou vectorielle. Les couches vectorielles peuvent être stockées dans la base de données PostgreSQL de MapX tandis que toute couche raster doit être appelée par WMS.

Les vues se basant sur des couches vectorielles peuvent être configurées directement dans l'interface de MapX via une fenêtre dans laquelle des paramètres comme les valeurs, les couleurs, l'opacité ainsi que la taille et la forme des symboles peuvent être définis (Figure 15). Les styles des vues raster ne peuvent pas être configurés dans MapX, ils doivent donc être déjà créés et associés à la couche dans GeoServer par exemple, avant que la vue soit appelée par WMS.

Value	Label	Color	Opacity	Size	Symbol
0	0		1	1	none
1	1 - 51		1	1	none
51	52 - 103		1	1	none
104	104 - 154		1	1	none
155	155 - 204		1	1	none
205	205 - 258		1	1	none

+ Rule x Last Rule x All

Reverse displayed layers order
If checked, the layer stack will be reversed : the top rule will draw a layer at the bottom of the stack

Missing values >

Ignore missing values
If checked, null or missing values will be ignored in the style

Figure 15: Fenêtre de configuration du style d'une vue dans MapX, ici la vue "Area of millet crop (hectares)" dans le projet "Risk - Niger".

Plusieurs vues basées sur des couches raster et vectorielles ont été créées dans MapX au sein de différents projets, à savoir "Risk – Niger" et "WESR: Climate". Ces données concernent notamment les émissions de gaz à effet de serre (CO₂ et CH₄) en 1970, 1990 et 2012, les anomalies de températures dans les centre urbains liées à l'activité humaine ou différentes simulations d'exposition aux risques naturels dans une région du Niger.

2.6.2. Méthode

Les vues des couches vectorielles pour les projets "Risk – Niger" et "WESR: Climate" ont été configurées manuellement dans MapX à partir de styles préalablement définis dans QGIS. Pour l'instant, la plateforme ne permet pas de définir un style automatiquement (à partir d'un fichier SLD par exemple). Le style doit donc être reproduit pour chaque intervalle/valeur souhaité. La Figure 16 ci-dessous montre la comparaison entre le style original dans QGIS et celui de la vue publiée dans MapX.

Pour les vues basées sur des couches raster, elles ont été créées dans MapX via WMS depuis le GeoServer du GRID – Genève. Les styles de ces couches ont été également configurés sur QGIS puis exportés sous forme de fichiers SLD, associés aux couches dans GeoServer. Le style de ces couches n'est donc pas configuré dans MapX mais est lié à GeoServer.

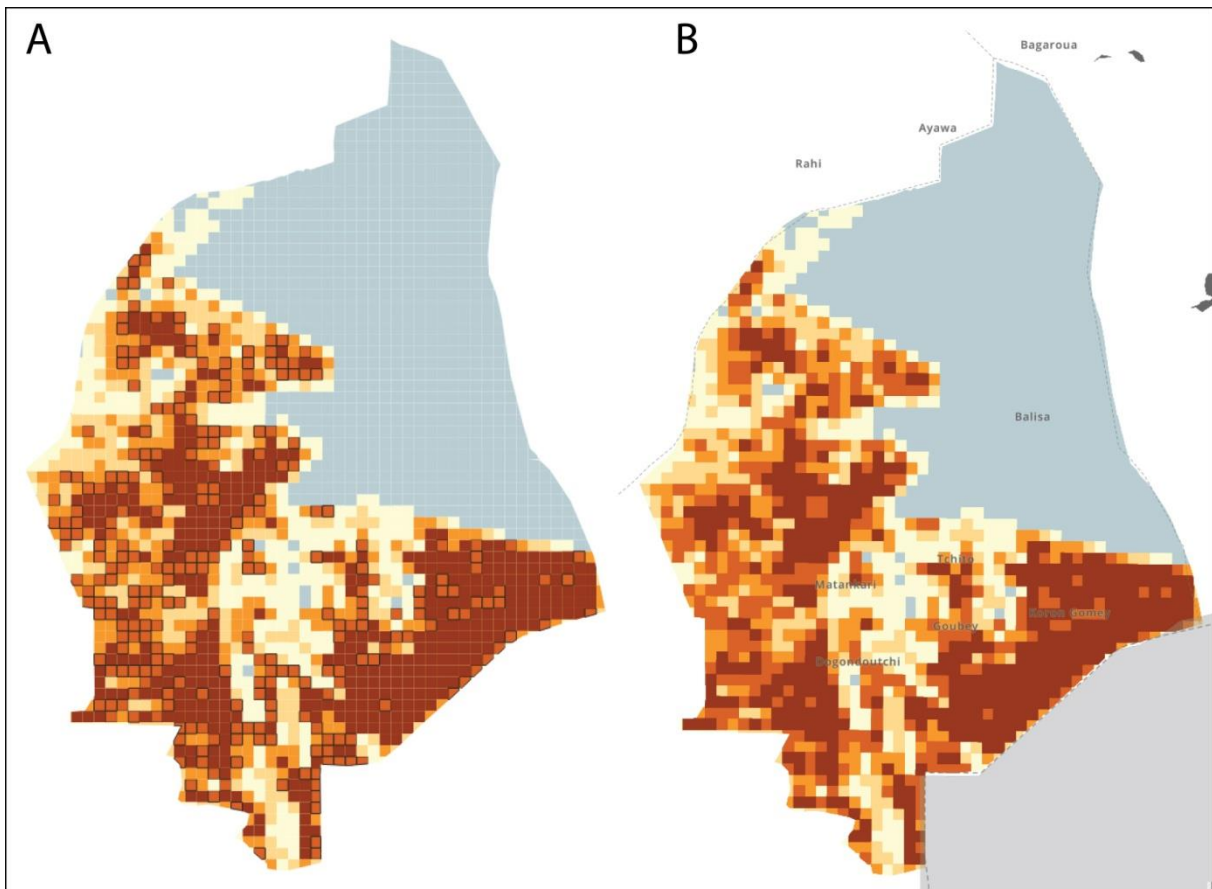


Figure 16: Comparaison entre les styles de la couche "Area of millet crop (hectares)" définis dans A) QGIS et B) MapX dans le projet "Risk – Niger".

2.6.3. Résultats

Dans le projet "Risk – Niger", 5 couches vectorielles et 12 couches raster ont été publiées, elles sont les suivantes:

- Area of millet crop (hectares)
- Average Annual Loss (USD)
- Average annual loss relative to its exposed value (unitless)
- Exposed value (USD)
- Subsoil USDA texture

- 12 Couches de "Drought disaster" calculées pour différentes périodes de retour de 25, 50, 100 et 250 ans, concernant respectivement la durée (en mois), l'intensité et la gravité (sans unités).

Dans le projet "WESR: Climate", 3 couches vectorielles et 8 couches raster ont été publiées, elles sont les suivantes:

- Trois couches présentant la vulnérabilité des pêcheries aux effets du changement climatique sur la période 2016 – 2050 "2016 2050 Vulnerability to Effects of Climate Change on Fisheries".
- Deux couches présentant la différence de température des centre urbains le jour et la nuit "Urban Heat Island UHI Summer Urban Rural Temperature Difference".
- 6 couches concernant les émissions mondiales de CO₂ et de CH₄ pour les années 1970, 1990 et 2012.

2.6.4. Bilan

La création de vues dans MapX a permis de se familiariser avec le processus de publication de données spatiales dans la plateforme. Cette mission a également été l'occasion de consacrer du temps aux tâches moins visibles liées à la création de vues comme le remplissage des métadonnées sur GeoNetwork, le processus de publication des données sur GeoServer et la gestion des fichiers de styles.

3. BILAN GÉNÉRAL DU STAGE

Ce stage au sein du GRID – Genève a tout d'abord été l'occasion de mettre en pratique et d'élargir les connaissances théoriques accumulées durant les cours du certificat complémentaire en géomatique de l'Université de Genève. Ces connaissances relèvent entre autres de traitement, d'utilisation et de visualisation de données spatiale, de la gestion de bases de données et de programmation informatique.

La nature variée des missions accomplies ainsi que leur impact concret ont assurément représenté un des aspects les plus positifs tout au long de ce stage de 3 mois et demi. Il a en effet été extrêmement gratifiant de collaborer à des projets dont la portée et les réalisations ont donné des résultats visibles et palpables.

Un autre aspect positif de ce stage a été de pouvoir évoluer au sein d'un service ayant un impact positif sur la gouvernance environnementale, que ce soit au niveau global, ciblé sur des pays ou même à une échelle plus locale. Il a été intéressant de voir les implications nécessaires à la production et au traitement de données qui sont par la suite employées par les décideurs politiques.

Collaborer au sein du projet MapX a particulièrement été intéressant de par le fait que cette plateforme est en constante évolution et qu'elle est amenée à se développer selon les besoins des utilisateurs et institutions clientes. Ce développement en fonction des besoins extérieurs amène également un rythme de travail non linéaire pour l'équipe de MapX, avec des périodes plus ou moins intenses.

4. RÉFÉRENCES

Articles:

Butler, H., Daly, M., Doyle, A., Gillies, S., Hagen, S., & Schaub, T. (2016). The geojson format (No. RFC 7946).

Weyns, Y., Hoex, L., & Matthysen, K. (2016). Analysis of the Interactive Map of Artisanal Mining Areas in Eastern DR Congo. 2015 Update. Antwerp: IPIS.

Sites web:

FEWS NET: page d'accueil, <http://fews.net/>

FEWS NET: Data Center, <http://fews.net/fews-data/333>

Geneva Water Hub: page d'accueil, <https://www.genevawaterhub.org/fr>

GRID – Genève, <http://www.grid.unep.ch/index.php?lang=fr>

IPIS: page d'accueil, <http://ipisresearch.be/>

IPIS: carte interactive des mines artisanales en RDC,
<http://www.ipisresearch.be/mapping/webmapping/drcongo/v5>

KoBoCollect: page d'accueil, <https://www.kobotoolbox.org/>

MapX: page d'accueil, <https://mapx.org>

MapX: plateforme, <https://app.mapx.org>

UN Environnement, <https://www.unenvironment.org/fr>

UNIGE: page du certificat complémentaire en géomatique <https://www.unige.ch/cgeom/>

5. REMERCIEMENTS

Pour m'avoir permis d'effectuer ce stage au sein du GRID – Genève, stage qui a été une expérience bénéfique au niveau professionnel mais également au niveau humain, je tiens à remercier le Dr. Pierre Lacroix, plus particulièrement pour sa disponibilité et son aide tout au long de ces quelques mois.

Je tiens également à remercier toute l'équipe du GRID – Genève, plus spécialement Thomas Piller et Antonio Benvenuti pour leurs conseils dans les nombreuses missions réalisées tout au long de ce stage.

6. ANNEXES

A) Script pour l'intégration des données KoBoCollect dans MapX

B) Script pour l'intégration des données FEWS NET dans MapX

```

1 #ANNEXE A) Script pour l'intégration des données KoBoCollect dans MapX
2
3 library("RPostgreSQL")
4 library("mailR")
5
6 #' Count the number of rows in a table stored in a database
7 #' @param con connection information
8 #' @param sourceId input table name
9
10 mxDbCountRow <- function(con,sourceId) {
11
12   # Test if source exists
13   sourceNotExists = isTRUE(!dbExistsTable(con,sourceId))
14
15   if(sourceNotExists) stop(sprintf(
16     "The source '%1$s' doesn't exist.",
17     sourceId)
18   )
19
20   # Count the number of rows
21   countRows <- dbGetQuery(con,
22     sprintf(
23       "SELECT count(*)
24       FROM %1$s;",
25       sourceId)
26   )
27   return(countRows)
28 }
29
30
31 #' Send an email to the administrator(s) in case of error
32 #' @param script_name input script name
33 #' @param recipients input list of recipients
34 #' @param cond input text returned by an error or a warning that will be sent by email
35
36 sendEmailError <- function(script_name,recipients,cond) {
37   send.mail(from = "bot@mapx.org",
38     to = recipients,
39     subject = sprintf("[Script failure] %1$s",
40       script_name),
41     body = sprintf("[%1$s] %2$s",
42       Sys.time(),
43       cond),
44     smtp = list(host.name = ****,
45       port = 587,
46       user.name = ****,
47       passwd = ****,
48       tls = TRUE),
49     authenticate = TRUE,
50     send = TRUE
51   )
52 }
53
54
55 #' Rename KoBo forms geolocation columns to "latitude" and "longitude"
56 #' @param df input dataframe - KoBo form
57
58 renameKoboLatLng <- function(df) {
59   colnames(df)[colnames(df) == "X_geolocation.0"] <- "latitude"
60   colnames(df)[colnames(df) == "X_geolocation.1"] <- "longitude"
61   return(df)
62 }
63
64
65 #' Convert a column in numeric format
66 #' @param df input dataframe - KoBo form
67 #' @param column input column
68
69 convertNum <- function(df,column) {
70   if((class(df[,column]) == "numeric" ) != "numeric") {
71     df[,column] <-as.numeric(df[,column])
72   }
73   return(df)

```



```

74 }
75
76
77 #' Delete rows without latitude and/or longitude
78 #' @param df input dataframe - clean KoBo form
79
80 deleteLatLngEmptyRows <- function(df){
81
82   # Checking for NA values in the longitude and latitude columns.
83   idx_na_lon <- (which(is.na(df [, "longitude"])))
84   idx_na_lat <- (which(is.na(df [, "latitude"])))
85   idx_tot_unique <- unique(c(idx_na_lon, idx_na_lat))
86
87   # Removing rows without values for latitude and longitude as they are not usable
   in a map
88   if(length(idx_tot_unique) > 0) {
89     df <- df[-c(idx_tot_unique),]
90   }
91
92   return(df)
93 }
94
95
96 #' Replace special characters by "_"
97 #' @param txt input text
98
99 replaceSpecialCharacters <- function(txt) {
100   gsub("[^a-zA-Z0-9]", "_", txt)
101 }
102
103
104 #' Remove file older than the specified number of days
105 #' @param outputDir input path to output directory
106 #' @param days input number of days
107
108 removeOldFiles <- function(outputDir, days){
109   date <- Sys.time() - (86400*days)
110   df <- file.info(list.files(outputDir, full.names = T))
111   old <- (rownames(df)[df$mtime < date])
112   file.remove(old)
113 }

```

```

1 # ANNEXE A) Script pour l'intégration des données KoBoCollect dans MapX (PARTIE 2)
2 # _____INSTALLING PACKAGES AND SETTING DIRECTORY_____
3
4 library(curl)
5 library(DBI)
6 library(geojsonio)
7 library(rgdal)
8 library(sp)
9
10 setwd("D:/MapX/kobo")
11
12 source("helper_functions.R")
13
14
15 # _____SETTING OPTIONS_____
16
17 opt <- list(
18   script_name = "KoBo to MapX: Haiti",
19   admins = c("thomas.piller@unepgrid.ch"),
20   username = ****,
21   password = ****,
22   form_id = ****,
23   url_prefix = "https://kc.humanitarianresponse.info/api/v1/data",
24   source_mapx = "mx_vector_spuhp_jm3io_ywgpx_drjur_u_17",
25   output_path = "./backups",
26   output_name = "kobo_compteurs",
27   output_minimum_storage_days = 30
28 )
29
30
31 # _____SETTING DB CONNECTION PARAMETERS_____
32
33 d <- list(
34   dbname = ****,
35   host = ****,
36   port = ****,
37   user = ****,
38   password = ****
39 )
40
41 dsn <- sprintf("PG:dbname=%s host=%s port=%s user=%s password=%s",
42               d$dbname,
43               d$host,
44               d$port,
45               d$user,
46               d$password
47 )
48
49
50 # _____SETTING CURL PARAMETERS_____
51
52 # Userpwd format: "username:password".
53 h <- curl::new_handle()
54 curl::handle_setopt(
55   handle = h,
56   httpauth = 1,
57   userpwd = sprintf("%1$s:%2$s",
58                     opt$username,
59                     opt$password)
60 )
61
62 # File format must be csv.
63 form_format <- "csv"
64
65
66 # _____IF NECESSARY, GETTING THE LIST OF ALL FORMS OWNED BY THE
67 USER_____
68
69 #If you don't know the id of the form you need to convert, uncomment and run the 3
70 # following lines before going further.
71 # url_forms_list <- sprintf("%1$s?format=%2$s",
72 #                             opt$url_prefix,
73 #                             form_format)

```

```

72 # forms_list <- read.csv(curl(url_forms_list, handle = h))
73 # View(forms_list)
74
75
76 # _____ DOWNLOADING DATA _____
77
78 url_form_dl <- sprintf("%1$s/%2$s?format=%3$s",
79                       opt$url_prefix,
80                       opt$form_id,
81                       form_format)
82
83 # Reading the form as a csv and storing it in a dataframe.
84 df <- tryCatch({
85   read.csv(curl(url_form_dl, handle = h))
86 },
87 warning = function(cond) {
88   sendEmailError(opt$script_name, opt$admins, cond)
89   stop(cond)
90 },
91 error = function(cond) {
92   sendEmailError(opt$script_name, opt$admins, cond)
93   stop(cond)
94 }
95 )
96
97
98 # _____ CLEANING DATA _____
99
100 # Renaming longitude and latitude columns.
101 df <- renameKoboLatLng(df)
102
103 # If necessary, converting longitude and latitude to numeric
104 df <- convertNum(df, "latitude")
105 df <- convertNum(df, "longitude")
106
107 # Deleting rows without latitude and/or longitude as they are not usable in a map
108 df <- deleteLatLngEmptyRows(df)
109
110
111 # _____ CREATING SPATIAL DATAFRAME _____
112
113 # Creating a spatial dataframe
114 spdf <- SpatialPointsDataFrame(coords = df[,c("longitude", "latitude")],
115                               df[,!(names(df) %in% c("longitude", "latitude"))],
116                               proj4string = CRS("+proj=longlat +datum=WGS84
117                               +ellps=WGS84 +towgs84=0,0,0")
118                               )
119
120 # _____ EXPORTING CLEAN DATA _____
121
122 #Setting the filename for output.
123 time <- replaceSpecialCharacters(Sys.time())
124
125 output_filename = sprintf("%1$s_%2$s",
126                            opt$output_name,
127                            time
128                            )
129
130 tryCatch({
131   geojson_write(spdf,
132                file = file.path(opt$output_path, paste0(output_filename,
133                  ".geojson")),
134                overwrite = TRUE,
135                encoding = "UTF-8")
136 },
137 warning = function(cond) {
138   sendEmailError(opt$script_name, opt$admins, cond)
139   stop(cond)
140 },
141 error = function(cond) {
142   sendEmailError(opt$script_name, opt$admins, cond)
143   stop(cond)

```

```
143 }
144 )
145
146 # _____ UPDATING MAPX DB _____
147
148 tryCatch({
149   writeOGR(spdf,
150     dsn = dsn,
151     layer = opt$source_mapx,
152     driver = 'PostgreSQL',
153     layer_options = c("FID=gid","GEOMETRY_NAME=geom"),
154     overwrite_layer = TRUE
155   )
156 },
157 warning = function(cond) {
158   sendEmailError(opt$script_name,opt$admins,cond)
159   stop(cond)
160 },
161 error = function(cond) {
162   sendEmailError(opt$script_name,opt$admins,cond)
163   stop(cond)
164 }
165 )
166
167 # _____ CLEANING BACKUPS _____
168
169
170
171 tryCatch({
172   removeOldFiles(opt$output_path,opt$output_minimum_storage_days)
173 },
174 warning = function(cond) {
175   sendEmailError(opt$script_name,opt$admins,cond)
176   stop(cond)
177 },
178 error = function(cond) {
179   sendEmailError(opt$script_name,opt$admins,cond)
180   stop(cond)
181 }
182 )
```

```

1 #ANNEXE B) Script pour l'intégration des données FEWS NET dans MapX
2
3 # _____INSTALLING PACKAGES & SETTING
  DIRECTORY_____
4
5 library(sf)
6 library(geojsonio)
7
8 setwd("ENTER DIRECTORY HERE")
9
10 # _____SETTING
   PARAMETERS_____
11
12 url_prefix <-
  "http://fews.net/data_portal_download/download?data_file_path=http%3A//shapefiles.fews
  .net.s3.amazonaws.com/HFIC/"
13
14 #Creating a list of all regions.
15 reg_names <- list(
16   "LAC",
17   "CA",
18   "EA",
19   "WA",
20   "SA")
21
22 #Creating a list of all urls for the download of all files.
23 reg_urls <- list(
24   url_reg_LAC = "LAC/caribbean-central-america",
25   url_reg_CA = "CA/central-asia",
26   url_reg_EA = "EA/east-africa",
27   url_reg_WA = "WA/west-africa",
28   url_reg_SA = "SA/southern-africa")
29
30 #List of destfiles in which data is downloaded.
31 destfiles <- list(
32   "LAC_data.zip",
33   "CA_data.zip",
34   "EA_data.zip",
35   "WA_data.zip",
36   "SA_data.zip")
37
38 #Setting current year and month
39 year <-format(Sys.Date(), "%Y")
40 month <- format(Sys.Date(), "%m")
41
42 #Setting the folder in which all files will be unzipped and two empty lists to be
  filled for the download process.
43 dsn_unzipped <- "global_data"
44 reg_files <- list()
45 tot_urls <- list()
46
47 # _____SETTING
   PARAMETERS_____
48
49 #Selecting the data version to be downloaded according to the current month.
50 #Download the appropriate file for each region.
51 #Unzip all files into one global folder.
52 if ((month == "03") | (month == "04") | (month == "05") | (month == "06")) {
53   latest_release = paste0(year, "02")
54   i <- 1
55   while (i <= length(reg_urls)) {
56     tot_url <- paste0(url_prefix, reg_urls[i], year, "02.zip")
57     tot_urls <- c(tot_urls, tot_url)
58     download.file(tot_urls[[i]], destfiles[[i]])
59     unzip(destfiles[[i]], exdir = dsn_unzipped)
60     i = i+1
61   }
62 } else if ((month == "07") | (month == "08") | (month == "09") | (month == "10")) {
63   latest_release = paste0(year, "06")
64   i <- 1
65   while (i <= length(reg_urls)) {
66     tot_url <- paste0(url_prefix, reg_urls[i], year, "06.zip")
67     tot_urls <- c(tot_urls, tot_url)

```

```

68     download.file(tot_urls[[i]], destfiles[[i]])
69     unzip(destfiles[[i]], exdir = dsn_unzipped)
70     i = i+1
71 }
72 } else if ((month == "11" | (month == "12"))) {
73     latest_release = paste0(year, "10")
74     i <- 1
75     while (i <= length(reg_urls)) {
76         tot_url <- paste0(url_prefix, reg_urls[i], year, "10.zip")
77         tot_urls <- c(tot_urls, tot_url)
78         download.file(tot_urls[[i]], destfiles[[i]])
79         unzip(destfiles[[i]], exdir = dsn_unzipped)
80         i = i+1
81     }
82 }
83
84 # _____READING & COMBINING
85 SHAPEFILES _____
86
87 reg_dis <- list()
88
89 #Reading each region shapefile, converting from "Multipolygon" to a collection of
90 separated polygons.
91 for (i in 1:length(reg_names)) {
92     reg_file <-paste0(reg_names[i], "_", latest_release, "_CS")
93     reg_files <- c(reg_files, reg_file)
94     reg_names[[i]] = st_read(dsn_unzipped, reg_files[[i]])
95     reg_dis[[i]] = st_cast(reg_names[[i]], "POLYGON")
96 }
97
98 #Merging all regions into one global layer.
99 FEWS_all_data <- rbind(reg_dis[[1]], reg_dis[[2]], reg_dis[[3]], reg_dis[[4]],
100 reg_dis[[5]])
101
102 # _____EXPORTING DATA AS
103 GEOJSON _____
104
105 FEWS_all_data_gj <- geojson_json(FEWS_all_data)
106 geojson_write(FEWS_all_data_gj, file ="FEWS_all_data.geojson")

```