



**HAL**  
open science

**Utilisation des données Google Street View pour  
cartographier la distribution géographique des espèces.  
Une étude préliminaire de la processionnaire du pin  
(*Thaumetopoea pityocampa*)**

Jérôme Rousselet, Charles-Edouard Imbert, Anissa Dekri, Jacques Garcia,  
Francis Goussard, Bruno Vincent, Olivier Denux, Christelle Robinet, Franck  
Dorkeld, Alain Roques, et al.

► **To cite this version:**

Jérôme Rousselet, Charles-Edouard Imbert, Anissa Dekri, Jacques Garcia, Francis Goussard, et al..  
Utilisation des données Google Street View pour cartographier la distribution géographique des es-  
pèces. Une étude préliminaire de la processionnaire du pin (*Thaumetopoea pityocampa*). Annales de  
l'INRGREF, 2012, 17, pp.173-188. hal-01189981

**HAL Id: hal-01189981**

**<https://hal.science/hal-01189981>**

Submitted on 1 Sep 2015

**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.

## Utilisation des données Google Street View pour cartographier la distribution géographique des espèces. Une étude préliminaire de la processionnaire du pin (*Thaumetopoea pityocampa*).

Jérôme Rousselet<sup>1</sup>, Charles-Edouard Imbert<sup>1</sup>, Anissa Dekri<sup>1</sup>, Jacques Garcia<sup>1</sup>, Francis Goussard<sup>1</sup>, Bruno Vincent<sup>1</sup>, Olivier Denux<sup>1</sup>, Christelle Robinet<sup>1</sup>, Franck Dorkeld<sup>2</sup>, Alain Roques<sup>1</sup> et Jean-Pierre Rossi<sup>2\*</sup>

<sup>1</sup>INRA, UR633 Zoologie Forestière, F- 45075 Orléans, France

<sup>2</sup>INRA, UMR CBGP (INRA/IRD/Cirad/Montpellier SupAgro), Campus International de Baillarguet, CS 30016, F-34988 Montferrier-sur-Lez Cedex, France

\*Auteur pour la correspondance

ملخص. استخدام بيانات قوقل ستريت فيو لرسم خريطة للتوزيع الجغرافي للأنواع. دراسة: PP: اولية لفراشة الصنوبر

(*Thaumetopoea pityocampa*)

ان رسم خرائط التوزيع الجغرافي للأنواع مهم جدا لرصد مجالات التوزيع الجغرافي للأنواع الاصلية والانواع الغازية. لسوء الحظ، قد تكون البيانات الضرورية صعبة المنال من الأدب وغالبية الثمن كي تجمع على أرض الواقع. لهذا السبب، قمنا بتصميم دراسة لاكتشاف كيف يمكن استخدام قواعد البيانات مثل جوجل ستريت فيو ( GSV ) لبيانات صالحة المثال. لقد اخترنا العمل على نوع من الحشرات، وهي فراشة الصنوبر *Thaumetopoea pityocampa* لان يرقات هذا النوع تنمو على أوراق الأشجار المضيفة وتنسج عشا ابيض يسهل رؤيته. وجود هذا النوع في موقع معين يمكن تحديده بسهولة من خلال دراسة المناظر البانوراما المتاحة للكثير من المواقع في قاعدة بيانات جوجل ستريت . وضعنا إجراءات موحدة لتقييم وجود PP من خلال بيانات GSV و قمنا بتنفيذها على مساحات تغطي 46848 كم<sup>2</sup> في وسط فرنسا. توزيع هذا النوع تم وصفه ايضا من خلال عينات اجريت على ارض الواقع. ان مقارنة البيانات الناتجة عن فحص صور جوجل ستريت فيو بالبيانات الميدانية بينت وجود فراشة الصنوبر بدقة 92.9% في منطقة دراستنا على شبكة 16 × 16 كم. وتشير هذه النتائج إلى أن استخدام قواعد البيانات GSV يمكن من تقديم بيانات ايكولوجية قيمة للأنواع التي يمكن تقديرها بصريا من الصور. هناك حاجة إلى مزيد من الدراسات لفهم أفضل لمجموعة من المقاييس المكانية من خلال النيات الدقيقة التي يوفرها GSV .  
فراشة لصنوبر/توزيع جغرافي/خرائط/قوقل ستريت فيو

**Résumé-** La cartographie de la distribution géographique des espèces est importante pour suivre l'évolution des aires de distribution d'espèces

invasives ou d'espèces natives en expansion géographique. Malheureusement, les données nécessaires sont parfois difficilement accessibles à partir de la littérature et sont coûteuses à collecter sur le terrain. Pour cette raison, nous avons conçu une étude dans le but d'explorer dans quelle mesure il est possible d'utiliser les bases de données telles que Google Street View (GSV) pour obtenir des données d'occurrence valides. Nous avons choisi de travailler avec une espèce d'insecte, la chenille processionnaire du pin (PP) *Thaumetopoea pityocampa* car les larves de cette espèce se développent dans le feuillage des arbres hôtes et tissent un nid blanc aisément visible. La présence de l'espèce dans un site donné peut donc être facilement renseignée en examinant les vues panoramiques disponibles pour de nombreuses localités dans la base de données de Google Street View. Nous avons conçu une procédure standardisée permettant d'évaluer la présence de la PP à partir des données GSV et nous l'avons mise en œuvre sur une aire d'étude couvrant 46 848 km<sup>2</sup> dans la région Centre en France. La distribution de l'espèce a également été décrite à l'aide d'échantillonnages réalisés sur le terrain. Les données issues de l'examen des images Google Street View ont été comparées aux données de terrain et se sont révélées de bons estimateurs de la présence de la processionnaire du pin avec une précision (proportion de valeurs correctement estimées) de 92.9% sur notre zone d'étude pour un maillage de 16 km x 16 km. Ces résultats suggèrent que l'exploitation des bases de données de GSV pourrait permettre de produire des données écologiques intéressantes pour les espèces dont la présence peut être estimée visuellement à partir de photographies. Des études complémentaires sont cependant nécessaires pour mieux cerner la gamme d'échelles spatiales auxquelles GSV fournit des données d'occurrence fiables.

### **Processionnaire du pin / distribution géographique / Cartographie / Google Streetview**

#### **Summary - Mapping species distribution using Google Street View. A preliminary study of the pine processionary.**

Mapping species distribution is an important and useful task to monitor invasive species spread or native species expansion under climate change. Unfortunately it requires a lot of occurrence data that are not easily available from literature and that are very time-consuming to collect in the field. For that reason, we designed a survey with the aim to explore to which extent large-scale databases such as Google Street View could be used to derive valid occurrence data. We worked with an insect species, the Pine Processionary Moth (PPM) *Thaumetopoea pityocampa* because the larvae of

that moth build silk nests on its host tree that are easily visible. The presence of the species at one location can therefore be inferred from visual records derived from the panoramic views available from Google Street View. We designed a standardized procedure allowing the evaluation of the presence of the PPM on a large sampling grid (covering 46 848 km<sup>2</sup>) located in France. Field sampling has been conducted in parallel, which allowed a straightforward comparison between field and Google-derived datasets. Data derived from Google Street View were highly similar to field data as we found an accuracy (percentage of field values correctly predicted using Google Street View) of 92.9%. We conclude that Google database might provide useful occurrence data for mapping the distribution of species which presence can be visually evaluated such as the PPM. More data are needed, however, to assess the range of spatial scales at which Google Street View actually provides reliable occurrence data.

### **Pine processionary moth / geographic distribution / mapping / Google street view**

#### **1. INTRODUCTION**

Aujourd'hui, un important effort de recherche est consacré à l'étude des aires de distribution d'espèces envahissantes ou d'espèces en expansion afin, notamment, de comprendre les modalités des invasions biologiques (Perrings *et al.*, 2010). Ce travail ainsi que la modélisation des aires de distribution demande d'importants volumes de données d'occurrence des espèces qui sont parfois difficiles à réunir à partir de la littérature ou à collecter sur le terrain. La question est d'autant plus problématique que l'on s'intéresse à des espèces dont la distribution évolue sous l'effet du changement climatique ou dans le cours d'une dynamique d'invasion. Durant la dernière décennie, les données géospatiales sont progressivement devenues plus accessibles avec l'apparition de ressources telles que Google Earth. Les ressources offertes par les services Google Earth ont ainsi été utilisées dans différents domaines de la recherche scientifique tels que la santé humaine ou animale (Carvalho *et al.*, 2012; Chang *et al.*, 2009), la biologie de la conservation (Asner *et al.*, 2009; Giri *et al.*, 2011) ou l'étude de la biodiversité (Guralnick *et al.*, 2007). Un niveau supplémentaire d'information spatiale a été atteint en 2007 lorsque Google a mis en service le système Google Street View (GSV)(Anguelov *et al.*, 2010). Ce système propose des vues panoramiques des abords des axes de communication correspondant à un niveau d'information local jamais atteint jusqu'alors. GSV n'est pas limité aux voies carrossables des villes et villages mais s'étend également dans les zones rurales de nombreux pays

(Anguelov *et al.*, 2010). Le système repose sur des véhicules, équipés d'un système de prise de vue géoréférencées, qui parcourent le monde et prennent des clichés qui sont ensuite gérés par Google et proposés en accès via l'internet (Anguelov *et al.*, 2010). Les vues panoramiques disponibles dans GSV permettent de se repérer dans le paysage, d'en identifier de très nombreux détails et d'explorer les routes, villes et villages.

Le travail présenté dans cet article a pour objectif d'explorer dans quelle mesure la technologie GSV peut être utilisée dans le domaine de la recherche scientifique afin de collecter des données sur la distribution géographique des espèces. L'espèce sélectionnée pour réaliser le travail est la processionnaire du pin (PP) (*Thaumetopoea pityocampa* Den. & Schiff., Lepidoptera, Notodontidae). La PP est un bon exemple d'espèce en expansion sous l'effet du changement climatique et offre plusieurs avantages vis-à-vis des objectifs de la présente étude. Il s'agit d'une espèce dont les larves consomment les aiguilles de différentes espèces des genres *Pinus* et *Cedrus* et notamment d'espèces largement répandues telles que *Pinus nigra*, utilisées comme essence forestière ou comme arbre ornemental en milieu urbain et rural. Les larves de PP tissent des nids d'hiver en soie de couleur blanche bien visible sur les conifères. La présence de nids indique la présence de l'espèce sans aucune ambiguïté car aucun autre organisme ne produit en hiver des structures comparables dans les pins et les cèdres rencontrés en Europe. Les nids de la PP sont agrégés dans l'espace (Arnaldo et Torres, 2005) et sont plus fréquents en lisière de peuplement ou le long des routes (Samalens et Rossi, 2011). Cette particularité est intéressante dans le contexte de notre étude qui vise à évaluer le potentiel informatif de vues panoramiques prises le long de voies de communication.

Nous avons, pour la première fois, testé l'utilisation des données disponibles dans le système GSV pour cartographier la distribution géographique d'une espèce et comparé ces résultats avec des données collectées sur le terrain afin d'explorer les limites de cette nouvelle approche.

## 2. MATERIEL ET METHODES

### 2.1. L'espèce modèle : la processionnaire du pin *Thaumetopoea pityocampa*

La processionnaire du pin est un insecte ravageur des conifères communs en Europe du sud et dans le pourtour du bassin méditerranéen où elle constitue le principal défoliateur des pins. La distribution géographique de cette espèce est en grande partie contrôlée par les températures minimales hivernales (Battisti *et al.*, 2005; Huchon et Démolin, 1970). Une étude

récente a montré que l'aire de distribution de la PP est en expansion vers le nord de la France et en altitude sous l'effet du changement climatique (Battisti *et al.*, 2005; Robinet *et al.*, 2007 ; Robinet *et al.*, 2010).

L'émergence des adultes se fait généralement durant l'été mais peut varier suivant les conditions climatiques locales. Peu après l'émergence et l'accouplement les femelles déposent leur ponte sur les aiguilles d'un arbre hôte. L'éclosion a lieu approximativement dans le mois qui suit, à savoir entre août et septembre en ce qui concerne la zone d'étude. Les larves sont grégaires, se nourrissent des aiguilles de l'arbre hôte et construisent un nid de soie (Démolin, 1969; Géri *et al.*, 1985). Les deux premiers stades larvaires tissent des pré-nids discrets qu'ils abandonnent régulièrement. A partir du 3<sup>ème</sup> stade larvaire, qui doit être atteint avant l'arrivée des premiers gels, les chenilles vont commencer à entretenir un nid dans lequel elles se développeront tout le reste de l'automne et de l'hiver. Ce nid est bien exposé au soleil et la soie nouvellement tissée à sa surface lui confère un caractère brillant par temps ensoleillé. La procession de nymphose se déroule entre la fin de l'hiver et le début du printemps suivant les conditions climatiques et météorologiques et les larves migrent dans le sol où elles achèvent leur développement. Les nids abandonnés passent d'une couleur blanche à du marron beige et deviennent rapidement moins perceptibles. Ils restent néanmoins observables jusqu'à la pousse du nouveau feuillage. La plupart des nids sont totalement dégradés avant la formation des nouveaux nids l'année suivante.

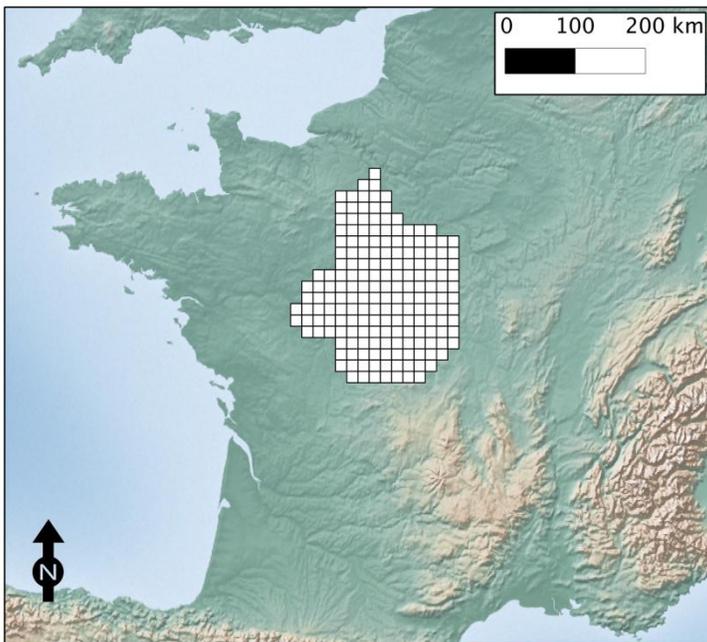
## 2.2. Fenêtre et grille d'échantillonnage

L'aire d'échantillonnage explorée dans ce travail représente une surface de 46 848 km<sup>2</sup> couvrant la région Centre de la France (figure 1). La surface d'étude a été discrétisée sous la forme de 183 cellules de 16 × 16 km représentant les unités d'échantillonnage (figure 1). Nous avons réalisé l'échantillonnage de la PP au niveau de chaque cellule suivant un protocole de mesure sur le terrain et une procédure basée sur l'exploitation des vues panoramiques disponibles dans la base de données GSV.

### 2.2.1 Protocoles d'échantillonnage sur le terrain

La PP a été échantillonnée suivant une approche dérivée d'un protocole standardisé utilisé pour suivre l'évolution de l'aire de distribution de l'espèce vers le nord de la France (Robinet *et al.*, 2010; Roques *et al.*, 2011). Dans le cadre de ce protocole « suivi du front », les cellules d'échantillonnage sont visitées par des observateurs circulant en voiture et recherchant la présence de nids d'hiver sur les espèces d'arbres hôtes.

Lorsqu'un nid est détecté, la cellule est considérée comme occupée, et les observateurs passent à une autre cellule. Lorsque toutes les routes et chemins carrossables ont été empruntés et qu'aucun nid n'a été observé, la cellule est considérée comme non occupée et donc la PP absente. Les observations douteuses, comme certains cas où l'arbre hôte est lointain sans pouvoir s'en rapprocher sont vérifiées par des observations à la jumelle. Les coordonnées des arbres hôtes sont enregistrées avec un GPS ou, dans le cas d'un arbre distant de la route, reportées sur un cliché photographique IGN géoréférencé dans les logiciels PhotoExploreur 3D (Bayo) ou Arpentgis mobile (D3E Electronique).



**Figure 1.** Aire d'étude et discrétisation en cellule d'échantillonnage. La zone d'étude a été discrétisée en 183 cellules de  $16 \times 16$  km (pour un total de  $46\,848$  km<sup>2</sup>). La processionnaire du pin a été échantillonnée dans chaque cellule et sa présence par ailleurs estimée par examen des vues panoramiques disponibles dans les bases de données GSV.

Les données de terrain utilisées dans cette étude proviennent de trois sources : (i) des données de présence / absence collectées par l'URZF dans

les cellules en limite de distribution selon le protocole « suivi du front » durant les hivers 2005-2006 et 2010-2011 (20 cellules sur 183) (ii) des données de présence récoltées par l'URZF en amont du front de colonisation selon un protocole similaire durant l'hiver 2007-2008 (157 cellules sur 183) (iii) des données de présence en amont du front de colonisation récoltées par le DSF (Département Santé des Forêts dépendant du Ministère Français de l'Agriculture, de l'Agroalimentaire et de la Forêt) (6 cellules sur 183).

### **2.2.2 Protocoles d'échantillonnage basés sur les ressources GSV**

L'ensemble des cellules d'échantillonnage a été exploré avec le système GSV et la présence de la PP évaluée avec un protocole dérivé du précédent (approche (i)). L'opérateur se déplace virtuellement le long des axes de communication répertoriés dans le système GSV en recherchant la présence de nid de PP sur les photographies panoramiques disponibles. Toutes les vues disponibles sont examinées.

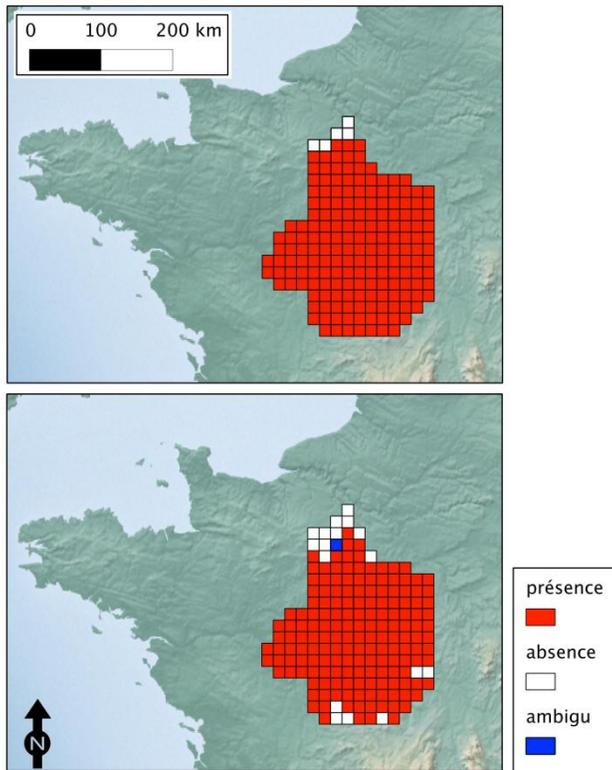
Le travail a été réalisé en septembre – octobre 2011 par un expérimentateur ignorant les données de terrain disponibles sur l'espèce. De façon similaire au protocole de terrain, l'opérateur ayant repéré un nid considère la cellule comme occupée et passe ensuite à une autre cellule. Lorsque l'identification d'un nid pose problème, l'opérateur fait varier le grandissement et utilise des angles de vue complémentaires afin d'améliorer son diagnostic. La position des arbres comportant un ou plusieurs nids ayant servi à attribuer le statut « occupé » à une cellule est géoréférencée avec Google Maps, Google Earth et/ou PhotoExploreur (Bayo-IGN). Dans chaque cellule, on observe également les différentes routes couvertes dans le dispositif GSV. Quatre types de situations ont été considérés : des cellules sans donnée GSV au moment de notre étude, des cellules avec et sans PP et des cellules pour lesquelles le statut de la PP reste ambigu. Ces dernières sont considérées par la suite comme des cellules non peuplées par la PP.

## **2.4 Analyses de données**

### **2.4.1 Comparaison des données de terrain et des données extraites de GSV**

Les jeux de données de terrain et les données tirées de l'analyse des images GSV se présentent sous la forme d'une grille de valeurs traduisant l'état des cellules d'échantillonnage (présence ou absence de l'espèce) (voir par exemple la figure 2). Nous avons abordé le traitement des données sous l'angle suivant : les données de terrain correspondent à une réalité dont les données issues de GSV constituent une prédiction. Le problème revient alors

à mesurer la performance d'un classificateur binaire (présence/absence) basée sur GSV. Nous avons calculé la sensibilité et la spécificité et différents indices afin de caractériser les performances de GSV en tant qu'estimateur des données de terrain.



**Figure 2:** Carte de distribution de la processionnaire du pin dans le centre de la France sur une grille à maille de  $16 \times 16$  km. En haut : observations réalisées sur le terrain En bas : observations issues de l'examen des vues panoramiques disponibles dans les bases de données GSV. Les cellules rouges représentent les zones où la présence de l'insecte a été détectée, les cellules blanches représentent les cas où l'insecte a été considéré absent, la cellule bleue correspond au cas ambigu pour lequel il n'a pas été possible de déterminer avec certitude si la PP était présente.

Pour cela, les cellules dans lesquelles la PP a été observée forment la classe des *positifs* et les cellules où la PP est absente forment la classe des *négatifs*. La comparaison entre la grille observée (donnée de terrain) et la grille issue des données Google a été réalisée en calculant le nombre de réalisations des 4 situations possibles (Fawcett, 2006) :

1. les vrais positifs (VP) : la PP est observée sur le terrain et dans les images Google
2. les vrais négatifs (VN) : la PP est absente des données de terrain et n'est pas observée dans les images Google
3. les faux positifs (FP) : la PP est absente d'après les données de terrain et présente sur les images Google
4. les faux négatifs (FN) : la PP est observée sur le terrain et absente sur les images disponibles dans Google

Les valeurs observées pour VP et VN ont été utilisées pour calculer la précision ou proportion de valeurs correctement prédites suivant la formule  $\text{précision} = (VP+VN)/(P+N)$  où P et N sont respectivement le nombre total de cas positifs et le nombre total de cas négatifs (Fawcett, 2006).

La sensibilité (ou taux de vrais positifs) mesure la proportion de cellule pour lesquelles la PP est présente (observation de terrain) correctement prédites à partir des images Google. La spécificité (ou taux de vrais négatifs) mesure la proportion de cellules où la PP est absente (mesure de terrain) bien prédites à partir des images Google. On estime la sensibilité comme le rapport VP/P et la spécificité comme FP/N (Fawcett, 2006).

Nous avons également calculé le coefficient de Matthews qui est une mesure directe de la covariation entre les observations issues de l'examen des images Google et les observations de terrain (Sing *et al.*, 2005). L'estimation est :

$$\text{Matthews} = (VP \times VN - FP \times FN) / ((VP+FN) \times (VN+FP) \times (VP+FP) \times (VN+FN))^{1/2}$$

Le coefficient de Matthews varie entre -1 et +1. Les calculs ont été réalisés avec le logiciel statistique R (R Development Core Team, 2012) et la librairie ROCR (Sing *et al.*, 2005).

### 3. RESULTATS

Les données collectées sur le terrain ont montré que la PP était présente sur une très large portion du territoire étudié : l'insecte a été observé sur 178 cellules et était absente de 5 cellules. Les observations dérivées de l'examen des images GSV diffèrent légèrement avec 165 présences, 17 absences et 1 cas classé comme « ambigu » car les images ne permettaient pas d'être

certain de la présence de l'insecte. Ce cas sera considéré comme une absence par la suite. Les résultats sont représentés dans la figure2.

**Tableau 1:** Matrice de confusion pour les occurrences de la processionnaire du pin observées au cours d'échantillonnage sur le terrain et à travers l'examen des images disponibles dans les bases de données de GSV.

		Observations de terrain	
		présence	absence
Observations basées sur l'examen des images GSV	présence	VP=165	FP=0
	absence	FN=13	VN=5

La confrontation entre les valeurs issues de Google et les données de terrain considérées comme référence montre une bonne correspondance entre les deux jeux de données. Sur les 183 cellules, 165 (90.2%) sont des Vrais Positifs (VP) c'est-à-dire que la PP a été observée effectivement sur le terrain et repérée sur les images GSV. Le nombre de Vrais Négatifs (VN) correspondant à une absence dans les deux cas s'élève à 5 (2.7%) (tableau 1). Nous avons observé 13 (7.1%) faux négatifs (FN) et aucun faux positifs (FP), cas où la PP n'aurait pas été repérée lors des études de terrain et observée à partir des images Google (tableau 1). Les valeurs obtenues pour VP, VN, FP et FN ont conduit à des valeurs de la sensibilité et de spécificité de 92,7% et 100% respectivement. Globalement, les données issues de l'examen des images Google sont de bons indicateurs des valeurs mesurées sur le terrain avec une précision de 92,9% et un coefficient de Matthews de 0,507.

#### 4. DISCUSSION

Les résultats présentés dans cet article montrent que les images disponibles dans les bases de données GSV peuvent fournir des informations très proches des résultats de campagnes d'échantillonnage réalisées sur le terrain. La précision, la sensibilité et la spécificité sont élevées. Cette observation est intéressante car elle ouvre des perspectives importantes pour simplifier l'acquisition de données ou la mise à jour d'informations préexistantes. Il faut cependant garder à l'esprit que l'adéquation entre les deux sources de données dépend en partie des échelles spatiales auxquelles nous avons opéré, en d'autre terme, rien n'indique qu'en travaillant à une autre échelle nous

aurions obtenu les mêmes résultats. L'échelle est définie par l'étendue (l'aire d'étude) et le grain ou résolution qui correspond à la taille des cellules utilisées dans notre étude. Le grain est ici analogue à la taille des unités d'échantillonnage dont on connaît l'impact sur la perception de la variabilité spatiale (Rossi et Nuutinen, 2004). L'augmentation du grain conduit à une diminution de la variabilité perçue (indiquée, par exemple par la variance du jeu de données). Dans notre cas, plus les cellules sont grandes, plus importante est la couverture par le système GSV et plus forte est la probabilité d'observer au moins un nid de PP si l'espèce est présente. Par ailleurs, des nids sur des arbres distants de la route sont plus difficilement discernables que des nids situés sur des arbres en bord de route. Sur le terrain, le doute peut être levé en utilisant des jumelles ou en empruntant une voie permettant de se rapprocher. La première solution est impossible dans GSV. D'autre part, plus la couverture GSV est importante, plus il est probable que les saisons auxquelles les clichés sont pris sont diversifiées à l'intérieur d'une cellule, et plus les chances d'observer un nid à un site et à une date favorables à sa détection sont grandes. Si l'on utilise des cellules plus petites, par exemple pour décrire plus finement le patron de distribution de l'espèce, on aura plus de chance de rencontrer des cellules avec peu ou pas de couverture par le système GSV et donc une mauvaise image de l'aire de distribution. Dans le cas de la PP, des carrés de  $16 \times 16$  km donnent des résultats très satisfaisants mais il serait intéressant de tester d'autres grains pour explorer les limites de validité du protocole présenté ici.

Les prises de vues qui alimentent la base de données GSV sont réalisées à différentes dates et, par exemple, un nombre important de cellules de  $16 \times 16$  km utilisées dans cet article correspondent à des clichés réalisés à au moins 2 dates différentes. C'est un avantage dans notre cas car ceci a limité le risque de ne disposer que de vues réalisées à une saison durant laquelle l'espèce étudiée n'est pas, ou est peu, détectable (en été principalement). Cette hétérogénéité entraîne cependant des effets difficiles à quantifier sur la variabilité spatiale de l'incertitude sur l'échantillonnage. Là encore, on peut supposer qu'il existe un fort effet « grain de l'étude » et ceci constitue une perspective intéressante pour des travaux complémentaires. La dimension temporelle des données hébergées dans le système GSV sera un élément clé dans l'avenir. En effet, les vues panoramiques seront mises à jour et il est probable que la couverture globale du système augmentera à l'avenir bien qu'aucune information ne soit aujourd'hui disponible. Dans une étude récente, Chen *et al.* (2011) ont montré que la comparaison de photographies prises sur un même site selon une même vue mais à un siècle d'écart permet

de décrire objectivement des modifications du paysage. Dans le cas d'espèces envahissantes, de telles séries temporelles d'observations pourraient permettre de documenter plus largement l'historique de processus d'invasion/expansion en disposant d'informations acquises selon un protocole reproductible et aux limites définissables. Des travaux récents en sociologie ont montré que GSV peut fournir des données permettant de caractériser l'environnement social en milieu urbain (Odgers *et al.*, 2012; Rundle *et al.*, 2011) et il est probable que de nouvelles applications verront le jour dans différents domaines de la science.

L'utilisation de GSV présuppose que l'espèce étudiée est visible sur des photographies ou qu'au moins certains indices indubitables de sa présence le sont. Dans le cas de la PP, les nids d'hiver sont très nettement visibles sur le fond vert des aiguilles des conifères mais de nombreuses espèces ne présentent pas d'indices aussi nets. L'agrégation des nids le long des lisières (Samalens et Rossi, 2011) et ou des axes de communication est également un aspect de la biologie de la PP favorable à son étude via GSV. Plus généralement, il semble possible d'appliquer le protocole présenté ici à des espèces de plantes ou d'arbres observables le long des axes routiers ou à des agents pathogènes provoquant des décolorations et des modifications de morphologie de la végétation. Dans la présente étude, on s'est intéressé à l'estimation de la présence de l'espèce sur une grille d'échantillonnage. L'estimation de la densité est une question beaucoup plus compliquée notamment parce qu'il devient nécessaire d'estimer la surface prospectée c'est-à-dire la surface sur laquelle on a recensé les individus (Southwood et Henderson, 2000). Déterminer si on peut parvenir à de telles estimations à partir des vues panoramiques hébergées dans les bases de données Google est une question ouverte et mériterait d'être examinée avec soin.

Si le terme même de « streetview » suggère que les villes sont les principales zones couvertes par le réseau, il ne faut pas pour autant conclure que les zones rurales ne sont pas cartographiées. Notre aire d'étude est vaste (figure 1) et nous avons pu vérifier au cours du travail que de très nombreuses routes de campagnes sont disponibles. A l'échelle de la Région Centre, ces routes traversent des paysages variés (zones urbaines, forestières, de grandes cultures ou de polyculture et de poly-élevage, ...) sans effet notable sur les résultats issus de la méthode GSV. Il est vrai cependant que la densité du réseau augmente aux abords des villes ce qui entraîne un certain niveau de variabilité spatiale dans l'incertitude liée à la mise en œuvre de la méthode. Cependant, nos résultats indiquent que ce facteur n'induit pas de fortes divergences entre les deux jeux de données. Dans le cadre des échanges commerciaux internationaux, les arbres à usage ornemental sont les premiers

pourvoyeurs d'insectes forestiers exotiques en Europe. La meilleure couverture des zones peuplées et urbanisées pourrait alors constituer un avantage si l'on considère que de nombreuses invasions biologiques pourraient démarrer depuis des villes et se propager ensuite vers les zones rurales (Smith *et al.*, 2007) en empruntant différents corridors (Säumel et Kowarik, 2010), dont notamment le réseau d'arbres ornementaux. La surveillance et l'analyse des phases initiales de dispersion des espèces envahissantes ainsi que l'identification des corridors pourraient être simplifiées par la densité de la couverture GSV dans les zones urbaines et périurbaines.

## 5. CONCLUSION

L'échelle à laquelle on travaille et en particulier le grain, ont probablement un rôle important sur l'adéquation entre données de terrain et données tirées de GSV. Dans le cas de la processionnaire du pin et à l'échelle spatiale de cette étude, l'exploitation des photographies disponibles dans la base de données GSV a permis d'obtenir des informations pertinentes sur l'aire de distribution géographique de l'espèce.

## 6. REMERCIEMENTS

Nous remercions le Département de la Santé des Forêts (DSF) et plus particulièrement François-Xavier Saintonge, Laurence Bouhot-Delduc, Hubert Pauly et Morgane Goudet. Ce travail a bénéficié du soutien financier de l'Agence National pour la Recherche (ANR) à travers le projet URTICLIM (Anticipation des effets du changement climatique sur l'impact écologique, sanitaire et social d'insectes forestiers urticants), du méta-programme INRA SMaCH et de la région Centre dans le cadre du projet ADRIEN (Arbres Dispersés et Rôle dans les Invasions d'Espèces Nuisibles). Nous remercions également le lecteur anonyme chargé d'évaluer ce travail dont les commentaires nous ont permis d'améliorer le présent article.

## RÉFÉRENCES BIBLIOGRAPHIQUES

Anguelov D., Dulong C., Filip D., Frueh C., Lafon S., Lyon R., Ogale A., Vincent L., Weaver J., 2010 - Google street view: Capturing the world at street level. *Computer* **43**, 32-38.

Arnaldo P.S., Torres L.M., 2005 - Spatial distribution and sampling of *Thaumetopoea pityocampa* (Den. & Schiff.) (Lep. Thaumetopoeidea) populations on *Pinus pinaster* in Montesinho, N. Portugal. *For. Ecol. Manage.* **210**, 1-7.

Asner G.P., Rudel T.K., Aide T.M., Defries R., Emerson R., 2009 - A contemporary assessment of change in humid tropical forests. *Conserv. Biol.* **23**, 1386-1395.

Battisti A., Stastny M., Netherer S., Robinet C., Schopf A., Roques A., Larsson S., 2005 - Expansion of geographic range in the pine processionary moth caused by increased winter temperatures. *Ecol. Appl.* **15**, 2084-2096.

Carvalho L.F.R., de Melo C.B., McManus C., Haddad J.P.A., 2012 - Use of satellite images for geographical localization of livestock holdings in Brazil. *Prev. Vet. Med.* **103**, 74-77.

Chang A.Y., Parrales M.E., Jimenez J., Sobieszczyk M.E., Hammer S.M., Copenhagen D.J., Kulkarni R.P., 2009 - Combining Google Earth and GIS mapping technologies in a dengue surveillance system for developing countries. *Int. J. Health Geogr.* **8**, 49.

Chen H., Yin K., Wang H., Zhong S., Wu N., Shi F., Zhu D., Zhu Q., Wang W., Ma Z., Fang X., Li W., Zhao P., Peng C., 2011 - Detecting one-hundred-year environmental changes in western China using seven-year repeat photography. *PLoS ONE* **6**.

Démolin G., 1969 - Comportement des adultes de *Thaumetopoea pityocampa* Schiff.: Dispersion spatiale, importance écologique. *Ann. Sci. Forest.* **26**, 81-102.

Fawcett T., 2006 - An introduction to ROC analysis. *Pattern Recog. Lett.* **27**, 861-874

Géri C., Millier C., Xeuxet D., 1985 - Mesure des populations de processionnaire du pin (*Thaumetopoea pityocampa* Schiff - LépidoptèreThaumetopoeidae) au Mont-Ventoux. *Ann. Sci. Forest.* **42**, 143-184.

Giri C., Ochieng E., Tieszen L.L., Zhu Z., Singh A., Loveland T., Masek J., Duke N., 2011 - Status and distribution of mangrove forests of the world using earth observation satellite data. *Global Ecol. Biogeogr.* **20**, 154-159.

Guralnick R.P., Hill A.W., Lane M., 2007 - Towards a collaborative, global infrastructure for biodiversity assessment. *Ecol. Lett.* **10**, 663-672.

Huchon H., Démolin G., 1970 - La bioécologie de la processionnaire du pin: Dispersion potentielle-dispersion actuelle. *Rev. Forest. Fr.* **22**, 220-233.

Odgers C.L., Caspi A., Bates C.J., Sampson R.J., Moffitt T.E., 2012 - Systematic social observation of children's neighborhoods using Google Street View: A reliable and cost-effective method. *J. Child Psychol. Psychiatry* **53**, 1009-1017.

Perrings C., Harold Mooney H., Williamson M., 2010 - *Bioinvasions and globalization. Ecology, economics, management, and Policy*. Oxford University Press.

R Development Core Team. 2012 - *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Robinet C., Baier P., Pennerstorfer J., Schopf A., Roques A., 2007 - Modelling the effects of climate change on the potential feeding activity of *Thaumetopoea pityocampa* (Den. & Schiff.) (Lep., Notodontidae) in France. *Global Ecol. Biogeogr.* **16**, 460-471.

Robinet C., Rousselet J., Imbert C.-E., Sauvard D., Garcia J., Goussard F., Roques A., 2010 - Le réchauffement climatique et le transport accidentel par l'homme responsables de l'expansion de la chenille processionnaire du pin. *Forêt Wallonne* **108**, 19-27.

Roques L., Soubeyrand S., Rousselet J., 2011 - A statistical-reaction-diffusion approach for analyzing expansion processes. *J. Theor. Biol.* **274**, 43-51.

Rossi J.-P., Nuutinen V., 2004 - The effect of sampling unit size on the perception of the spatial pattern of earthworm (*Lumbricus terrestris* L.) middens. *Appl. Soil Ecol.* **27**, 189-196.

Rundle A.G., Bader M.D.M., Richards C.A., Neckerman K.M., Teitler J.O., 2011 - Using Google Street View to audit neighborhood environments. *Am. J. Prev. Med.* **40**, 94-100.

Samalens J.C., Rossi J.-P., 2011 - Does landscape composition alter the spatiotemporal distribution of the pine processionary moth in a pine plantation forest? *Popul. Ecol.* **53**, 287-296.

Säumel I., Kowarik I., 2010 - Urban rivers as dispersal corridors for primarily wind-dispersed invasive tree species. *Landscape Urban Plann.* **94**, 244-249.

Sing T., Sander O., Beerenwinkel N., Lengauer T., 2005 - ROCR: Visualizing classifier performance in R. *Bioinformatics* **21**, 3940-3941.

Smith R.M., Baker R.H.A., Malumphy C.P., Hockland S., Hammon R.P., Ostojá-Starzewski J.C., Collins D.W., 2007 - Recent non-native invertebrate plant pest establishments in Great Britain: Origins, pathways, and trends. *Agric. For. Entomol.* **9**, 307-326.

Southwood T.R.E., Henderson P.A., 2000 - *Ecological methods*. Blackwell Science.