Landspotting – Games for Improving Global Land Cover

Tobias Sturn
Institute of Computer Graphics and Algorithms
Vienna University of Technology, Austria
sturn@cg.tuwien.ac.at

Michael Wimmer
Institute of Computer Graphics and Algorithms
Vienna University of Technology, Austria
wimmer@cg.tuwien.ac.a

Peter Purgathofer
Institute for Design and Assessment of Technology
Vienna University of Technology, Austria
purg@igw.tuwien.ac.at

Steffen Fritz
IIASA
fritz@iiasa.ac.at

ABSTRACT
Current satellite-derived land cover products, which are very important for answering many crucial questions, show huge disagreements. In this paper, we introduce four serious game prototypes – a Facebook strategy game played on Google Maps, a Facebook tagging game, a tower-defense game, and an aesthetic tile game for the iPad – with the purpose of improving global land cover data. We describe the games in detail and discuss the design decisions we made and challenges we faced while developing the games. We evaluate how much the players have already been able to improve global land cover data and provide evidence that games can be a useful way to increase the quality of this data. Finally, we discuss how the main game is being perceived by the players and what has to be further improved to attract a bigger audience.

Categories and Subject Descriptors
D.2.2 [**Software Engineering**]: Design Tools and Techniques; K.8.0 [**Personal computing**]: General-*Games*

Keywords
Serious Games, Global Land Cover, Crowd-Sourcing, Landspotting, Geo-Wiki

1. INTRODUCTION
At present, there is no single satellite-derived global land-cover product available that is accurate enough to provide reliable information to answer very important questions for humanity like "Does the climate change?" or "How much additional land is available to grow biofuels?", or to tackle problems of food security [15].

In the last decade, three global satellite-derived land cover products have been created: MODIS[1], GlobCover[2], and GLC-2000[3]. A pixel-by-pixel comparison of these three different land-cover datasets reveals huge areas of the world where they do not agree [15]. For example the three products have a forest and cropland disagreement of 893 million hectare and estimates of cropland differ by up to 20% which makes these maps very unreliable [8]. To improve these land-cover products, the Geo-Wiki [7] has been introduced to crowd-source user validations to create a new and better hybrid map. Unfortunately, it is quite difficult to gather enough validations without giving the users any incentives to validate land.

Therefore, the goal of the Landspotting project[4] is to use games to crowd-source land-cover validations. The incentives for the players to contribute should be the same as for people to play games, namely fun. Using games to improve global land-cover datasets has not been done before and is an exciting new way how games can be used for a serious purpose.

Figure 1 shows the games developed so far. The main game for improving land-cover data (top left, logo) is a Facebook real-time strategy game played on Google Maps. The second game (top right) is a Facebook tagging game with which we try to find out what annotations players use for land cover. The third game (bottom right) is an iPad tower defense game. The last game is an iPad tile game which uses the aesthetics of satellite data.

In this paper, we describe in detail the games and the decisions we made and the challenges we faced while developing the games for improving global land cover and the insights we gained. We explain how we try to make the games fun and engaging and how we give the player incentives to categorize the earth correctly. An evaluation is given how much the land-cover data has been improved through the games. We finally present a small study how our main Landspotting game is perceived by the players and what has to be improved to attract a bigger audience.

[1] http://modis.gsfc.nasa.gov
[2] http://ionia1.esrin.esa.int
[3] http://bioval.jrc.ec.europa.eu/products/glc2000/glc2000.php
[4] http://www.landspotting.org

Figure 1: The Landspotting games

The main *contributions* of this work are:

- Exploring different platforms for games with a scientific purpose.
- Extensive use of *social* game mechanics for games with a purpose.
- Examples of disjoint human computation game design for improving global land cover.

2. RELATED WORK

The idea of using games for a scientific purpose (GWAP, games with a purpose) is not new, and some games have already been able to very impressively help science. One of the most famous examples is *FoldIt* [6, 5], a puzzle-game in which the players fold proteins. Within ten days, the FoldIt players have, for example, accurately approximated the crystal structure of M-PMV, an AIDS-causing simian virus, which has been an unsolved problem for scientists for the last 15 years. Just recently, with the help of the 230, 000 FoldIt players, a novel algorithm for protein folding has been found which outperforms previously published methods [9].

Several types of game mechanics have been developed to gather correct meta-information through games. One group of GWAPs are the so-called *output agreement* games, for example the *ESP* game[5] [1]. In this game, two random players have to label a given image (the input). The more they agree on the labels (the output), the higher scores they get – hence output agreement. As of July 2008, 200, 000 players had contributed to more than 50 million labels.

Another group of GWAPs are described as *input agreement* games, for example *TagATune*[5] [12]. In TagATune, a sound clip (the input) is presented to two random players, who in return give as output a series of labels to the other player. The two players win the game if they both correctly agree whether they have heard the same input sound or not.

[5] http://www.gwap.com

A third group of GWAPs are called *inversion problem* games, for example the game *Verbosity*[5] [2]. One player, the describer, is given a word as input and has to describe this word to the guesser. The two players only get points if the output of the guesser represents the input of the describer. For example, the describer describes the word "milk" with "white, something to drink, people usually eat cereal with it". Thus, facts are collected as a side effect of playing.

In the image annotation game *Kisskissban* [4], a third player, the blocker, is added, who competes with two other collaborative players, the couple. While the couple tries to find consensual descriptions of an image, the blocker's mission is to prevent the couple from reaching consensus. The blocker will try to detect and prevent coalition between the couple. Therefore, these efforts naturally form a player-level cheating-proof mechanism. To evade the restrictions set by the blocker, the couple would endeavor to bring up a more diverse set of image annotations.

As can be seen from the games described above, there are already a number of GWAPs with impressive scientific results. Also, several types of mechanics for collecting correct meta-information through games have been developed. But there are still no serious games which have achieved the tremendous success of games like Farm Ville, World of Warcraft or Angry Birds, with hundreds of millions of players per month. As has already been described by Markus Krause [11, 10], the currently tightly coupled GWAPs do not use the whole scope of game-design elements. Instead, he proposes a disjoint game design where the task to be solved should be part of the mechanics, but no longer the dominant element. Therefore, our goal in the Landspotting project is to take successful games, especially social games like FarmVille, as reference and to try to copy these already proven game mechanics to create games with a scientific purpose who appeal to a large audience.

In order to be able to get a first impression how much fun the developed games are, they are compared against the ten high-level heuristics by Paavilainen [14], which define what makes a social game entertaining, listed in Table 1.

Spontaneity	Easily accessible and understandable.
Interruptability	Short game rounds shall be supported.
Continuity	Incentives for coming back. Feel progress.
Discovery	Achievements and new content regularly.
Virality	Virality with invite messages.
Narrativity	Broadcast game events to arouse curiosity.
Sharing	Honor sharing resources and information.
Expression	Opportunity to express themselves.
Sociability	Honor communication with friends.
Ranking	High score lists motivate to compete.

Table 1: High-level heuristics for social games presented by Paavilainen [14].

3. GAME DESIGN

This chapter describes the design process and the design decisions made for how to get correct land-cover data and how to make the Landspotting games fun and appealing to reach high player numbers. We have decided to create a number of games and pick out of these games the most promising ones on which to concentrate further. Therefore, we have developed four game prototypes, which will be discussed in more detail in the next chapter. Our team consists of 2 professors, 2 project assistants and 5 students, which imposes restrictions on graphics, 3d models, sound effects, storyboard and character design, thus the explored game ideas have to fit within these resource restrictions.

3.1 Disjoint Game Design

The first decision was to use a *disjoint game design* [11], i.e., one where the actual task of "landspotting" the earth is only part of the game mechanics and is no longer the dominant element. We try to achieve this by emulating already successful games and by trying to add the additional "landspotting" task as smoothly as possible to these already proven games. For example, one game tries to copy the game mechanics of the very successful game series Civilization. Another game tries to emulate the proven tower-defense mechanics like in Plants vs. Zombies.[6] Another game uses similar mechanics as the ESP game [1], but with many social elements.

3.2 For Which Platforms?

We analyzed several platforms and determined how feasible they are for a Landspotting game prototype. Game consoles were discarded due to lack of the required development kits. We have taken the following platforms into consideration: PC standalone application, standalone browser games, Facebook browser games, iPad, iPhone, and Android phones.

We found that the Facebook platform would be the most promising option for several reasons. First, the player numbers on Facebook are still very impressive (even though they have dropped a bit in the last couple of months). The number-one game, FarmVille 2, is still played by 57 million people per month. Second, the assets needed for a Facebook game are in general less than for a PC standalone applications. Third, the social aspect, which is the foundation of most Facebook games, is very important in the Landspotting project. Finally, the fact that Facebook games are playable over the browser without having to install the application or demanding high hardware specifications is very important to gain a large audience.

Another platform which also turns out to be very interesting are mobile devices like the iPhone, iPad, or Android devices. In the last couple of years, the use of smart phones has increased tremendously. For example, the mobile game Draw Something[7] peaked at over 15 million users per day. Therefore, we have also created games for these mobile platforms and have also investigated the possibilities of HTML5 [3] to make cross-platform games.

3.3 Problem Space

One property of land-cover data is that the problem space is not unlimited, and does not change quickly. The Earth's surface consists of 149 million square kilometers of land and 361 million square kilometers of water. Only land has to be validated by the Landspotting games. For practical land-cover validation, the surface is partitioned into equally sized tiles, each of which receives a unique land-cover type. Table 2 shows how many players are needed at different tile sizes to land-cover the whole earth. At a tile size of 1 square kilometer, and assuming one player validates on average 100 tiles, 1.48 million players are needed for this task. This number looks huge, but when comparing this number with the player numbers of games mentioned in the last chapter they seem more accessible.

[6] www.popcap.com/games/plants-vs-zombies

[7] www.omgpop.com/drawsomething

Tile area	Nr of tiles	Players needed
0,1 km^2 (100 m^2)	1,490,000,000,000	14,900,000,000
0,5 km^2 (500 m^2)	298,000,000	2,980,000
1 km^2	149,000,000	1,490,000
10 km^2	14,900,000	149,000
100 km^2	1,490,000	14,900

Table 2: How many players are needed? (With 1 player exploring 100 tiles)

3.4 Ensuring Correctness of Data

One of the big challenges for a GWAP is getting correct data from the players, since gamers will be looking for exploits and the limitations of the game mechanics. One important aspect is that the main incentive of the player to play the game is to have fun, so providing correct data should be the most rewarding strategy for a player in that respect.

We have analyzed the different approaches to guarantee data correctness described in Section 2. The *inversion problem* cannot be used for Landspotting because, in most cases, it is not possible for the guesser to guess the exact position of the map from land-cover descriptions alone. Therefore it is not possible to check if the describer and guesser see the same map. *Input agreement* is also very difficult to apply because it is impossible to decide if the land-cover maps have the same position based only on land-cover descriptions. Therefore, we use *output agreement*. The more the players agree on the land cover they have given as output, the higher scores they get and the faster they progress in the game. We try to use presents and achievements as positive reinforcements to educate the players to land-cover the earth correctly. Beside ensuring data correctness, we try to collect as detailed land-cover information from the players as possible. For example, in one game we let players annotate the land with simple drawing tools like a brush instead of using a simple classification. Thus, we gain very detailed land-cover maps, containing more information than simple annotations.

3.5 2D vs 3D

In a series of early prototypes, a number of 3D games were developed. One example is an archeology game in which the player has to find hidden treasures like dinosaur bones or diamonds in a 3D scene with different tools like a pick or a shovel. Depending on the tools used, the land-cover data could be estimated. It soon turned out that 3D is not good for estimating land-cover data. Most terrain, especially those where most disagreements in the datasets occur (e.g., Africa), is rather flat, and therefore the elevation information is not important. Also, close to the ground very good texture quality is needed, otherwise the texture cannot be

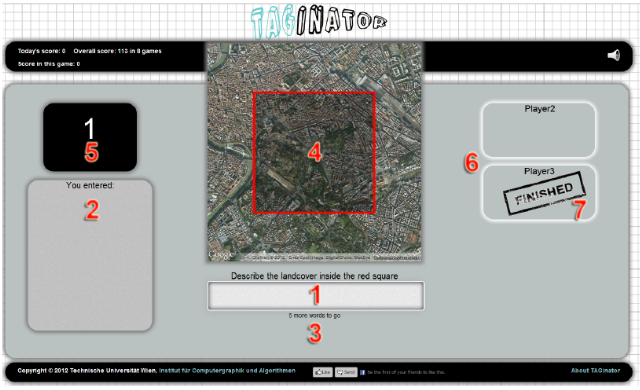

Figure 2: Annotating land cover in TAGinator.

recognized anymore because of the high magnification. Finally, much more time is needed for creating assets like 3D models. Therefore, we decided to concentrate on 2D games instead. This also makes sense for a Facebook or mobile game, where the majority of games are in 2D.

4. THE GAMES

This section describes the four developed game prototypes: a tagging game, a strategy game, a tower-defense game, and an aesthetic tile game.

4.1 TAGinator

Game Mechanics. The first Landspotting game prototype is TAGinator (see Figure 2).[8] It is a social Facebook game with the aim of collecting annotations for land-cover maps. In contrast to the other games, the main goal was not only to get land-cover classifications according to predefined categories, but to investigate what categories users would choose if they are given the choice to freely annotate a tile. This allows verifying whether the land-cover classifications proposed by the other games and by Geo-Wiki are intuitive. The game is implemented in HTML 5 [3] and can therefore be played on any platform with a browser, including PC, Mac, iPhone, iPad, and Android.

Ensuring Correctness of Data. TAGinator is an *output agreement* game and is quite similar to the ESP game. Instead of images, land-cover maps (4) are shown to the player, who has to enter (5) as many annotations as possible about the land-cover map into the text area (1) under time pressure. The more he agrees with the other players (6, 7) on the annotations, the higher score he gets. Through this mechanic, correctness is guaranteed.

Aesthetics. Many incentives are provided for the players to play the game regularly and to fulfill the requirements of what makes social games entertaining (Table 1 [14]). As heuristics like "Spontaneity", "Interruptability" and "Continuity" are already realized through the game design itself, we decided to implement the heuristics "Discovery", "Sociability" and "Ranking", too. The player's scores are presented per game round, per day and overall. By having separate scores, it is possible to present high-score lists for the all-time bests, monthly bests or daily bests. That way the possibility to climb the ladder and to compete with other players by achieving higher scores always becomes the focus of attention and motivates players to spend their time in the game. Beyond that, achievements can be collected as a further motivation, and players can post their scores on their Facebook wall as well as invite friends to play the game. High-score lists are shown on the starting page as well as achievements already collected and those still receivable. Communication channels like chats were omitted intentionally, and the names of fellow players are not shown during the game to prevent users from trivially agreeing on cheating strategies. To still fulfill the requirements of the heuristic of sociability, users are able to post messages on their Facebook wall or invite friends to play the game. Further on, at the very end of the game, when the final scoring round is finished, the names of the fellow players are shown to amplify the competitive character of the game. Additionally, sound effects were added to provide a more immersive gameplay experience.

4.2 Strategy Game

Game Mechanics. The main Landspotting game prototype[9] is a multiplayer online real-time strategy Facebook game played on Google Maps (see Figure 3). The game was implemented in Flash, like most Facebook games, which restricts the game to platforms with Flash support. The basic idea for this game is to use the successful, proven game mechanics of the famous strategy game Civilization[10] and to harmonically add the task "Landspotting" to the game. In Civilization, the players try to conquer the world by ruling and building up their own empires. The players build cities, harvest the land within and around their cities, build new troops to either defend their cities or to attack other empires, research new technologies, and make alliances with other players. The players need to think strategically and financially in order to plan and build a wealthy, prospering empire. This whole setting is perfect for the task "Landspotting". The big difference of the Landspotting game is that instead of playing in an artificial scenario like in Civilization, the players play on the actual Earth. The way we integrated Landspotting was via the idea of an "exploration task": before a unit can be moved to a new position, the area has to be explored by annotating the land. The player draws on a map of the new area (tile) using standard image editing tools (brush, fill bucket), see Figure 4. The user can also change the transparency of the land cover, undo and redo, and use key shortcuts to make drawing more comfortable.

Ensuring Correctness of Data. To guarantee data correctness, we again use *output agreement*, like in TAGinator. The more the players agree on the land cover they have drawn, the higher scores they get and the faster they progress in the game. After each good exploration, the player receives a "present", for example with a message like "Congratulations! You have made a great exploration! As a present you can now build cities!". In order to make output agreement work, we have to make sure that the players land-cover the same maps. Two mechanics have been added to guaranteeing this:

[8] https://apps.facebook.com/taginator

[9] https://apps.facebook.com/landspotting

[10] http://www.civilization.com

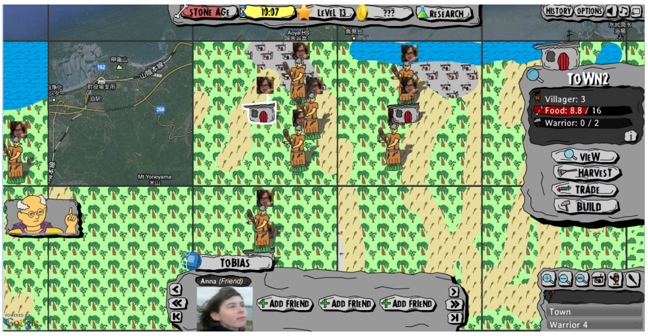

Figure 3: The Landspotting strategy game. Units, cities and the explored land are depicted. On the top is the status bar with information about the current state of the game. On the left is the advisor who proposes tasks to do next. On the bottom is the friends bar, like in other social Facebook games.

Figure 4: Exploring land in the strategy game. The brush size, type and transparency can be chosen in the right window. Undo, redo and fill are also supported.

- New players always start at the newest city built by another player. This guarantees that the land where the new player starts has already been explored. Therefore we can teach the player from the beginning that he has to comply with other players to progress faster. This mechanic also has the advantage that all players are close to each other, which is necessary for a strategy game, where players are supposed to compete. The game will automatically expand, because new players are set on the border of explored land.
- In Facebook games like FarmVille, it is very common to pay real money to rush the time to finish building something faster. We also use this *rush* mechanic, but instead of letting the player pay real money, we let the player "pay" by land-covering land which has already been explored by another player. The more he agrees with the other player, the more time is rushed and the faster the player progresses in the game.

Aesthetics. To keep the players inside the game, we have used many of the already proven social-game mechanics listed in Table 1. For example, we have added an advisor who helps the player at the beginning of the game (to fulfill the heuristic of "Spontaneity" to make the game easily accessible), and who proposes tasks throughout the game that the player could do next to keep the player engaged (to implement the heuristic of "Continuity" that the player feels progress and that he has incentives for coming back later). In addition, common Facebook game mechanics like sending private messages and posting wall messages have been added to meet the requirements of "Narrativity" to broadcast game events to arouse curiosity and of "Sociability" to let the players communicate with each other. It is also possible to invite friends to the game to fulfill the heuristic of "Virality". Players can also send each other gifts like giving a unit as present to fulfill the heuristic "Sharing". Furthermore, the game maintains a global high-score list and provides achievements to satisfy the needs of the competitive players and the heuristics "Ranking" and "Discovery". Sound effects and music are used to make the game more immersive. For example, if the player marks an area with tree cover, birds begin to sing, if the player attacks another empire with a unit, fighting sounds are played, and after the player has researched music, he can listen to drum beats.

4.3 Tower-Defense Game

Game Mechanics. The third Landspotting game prototype is an iPad tower-defense game which emulates the very successful game Plants vs. Zombies, which has sold more than 300,000 copies in the first nine days. It can be seen as a stripped-down version of the strategy game described in Section 4.1. The player has to defend his city from quickly approaching enemies by building warriors, towers, walls, and traps to stop these threats from reaching his city. The Landspotting task is added to the game similarly as has been done in the strategy game. To speed up the time needed for building something, the player has to gather resources by exploring foreign land. The player does this by picking a land-cover map, randomly taken from the strategy game, and land-covering this map by drawing the land cover with a brush. For drawing these land-cover maps, the iPad is the perfect tool because it feels much more natural to draw with the fingers than with the mouse.

Ensuring Correctness of Data. To ensure correct data, again output agreement is used. The more the players agrees with the land covers of other players, the more time is rushed and therefore the faster his units and objects get built. This is important because after some time, more and more dangerous enemies appear, and therefore more and more units and buildings are needed to defend the city.

Aesthetics. To motivate the players to play the game to compete with their friends and to fulfill the heuristics "Ranking", "Continuity," and "Discovery," Game Center is supported in the game, which provides a leaderboard and achievements to unlock. The game also integrates Facebook and Twitter support to satisfy the heuristics "Virality," "Narrativity," and "Sharing". Furthermore, sound effects and music have been added to make the game more immersive.

4.4 Tile World

Game Mechanics. The fourth game prototype is an aesthetic tile game that relies on the aesthetics of satellite im-

Figure 5: Tile World: an aesthetic tile game.

Urban	city, urban, town, houses, streets
Tree	trees, woods, forest
Grass	grass, grassland, meadows
Shrub	shrubs, bushes
Cult. Soil	cult. soil, fields, soil, agriculture, farmland
Uncult. Soil	uncultivated soil
Stone	rocks, mountain, hills, stone, ridges
Desert	desert, sand, dunes
Water	water, river, lake, sea, ocean
Snow/Ice	snow, ice, icefield
Unknown	unknown, undefined, nothing, darkness

Table 3: Annotations collected by TAGinator.

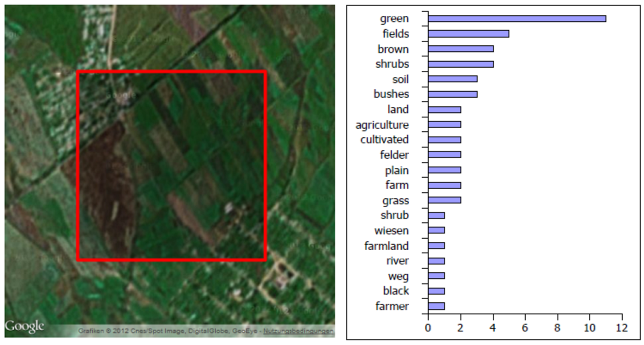

Figure 6: Annotations for a map with bad quality, with a plot of frequencies.

ages (Figure 5). The idea is that tiles of different land covers placed in a grid can form a nice images. A grid with empty tiles of different categories is presented on the left (1). The player can choose from the right list (2) tiles of the correct type and drag the tile to the correct position on the grid. As a result, the player has created a simple grid, which can later be combined with other simple grids to create a bigger, interesting mosaic.

Ensuring Correctness of Data. We have intentionally not used any mechanics for preserving correct data in Tile World. The incentives for the players to cheat are not high, as they do not compete against other players. In a later step, the correct land-cover type is calculated based on *output agreement*. The more the players agree on the land-cover type of a tile, the more the game assumes the tile to be of that type.

Aesthetics. To fulfill the heuristics "Sociability" and "Virality," the user can save the mosaics either to his device or share them with friends on social platforms like Twitter and Facebook. The game is very easy to understand and accessible, and is perfect for short game rounds. Therefore, the heuristics "Spontaneity", "Interruptability," and "Expression" are already implemented through the game design itself. Finally, a high-score list for who has made the biggest mosaic and daily achievements have been added to meet the heuristic of "Continuity", "Discovery," and "Ranking". The game has been implemented for the iPad and will eventually be available on the AppStore.

5. RESULTS

In this section, we analyze the results players generated with the prototypes and how much our main Landspotting game has already been able to contribute to improving global land-cover data.

5.1 Collected Annotations

With the TAGinator game, which has been played by 38 players, 247 distinct annotations out of 1,280 land-cover annotations from 20 different land-cover maps (each 2km^2 in size) with varying resolutions and qualities have been collected so far. Table 3 shows the most frequently given annotations, ordered by the number of occurrences and type. Note that players were meant to type in their own annotations, as discussed earlier.

In Figure 6, one of the annotated land-cover maps (42.3309 latitude, 77.9236 longitude) of very low quality can be seen with the top 20 annotations given by the players. While not all top 5 annotations contain meaningful information, such as "green" and "brown", the annotations "fields", "shrubs" and "soil" are very useful land-cover annotations. These player annotations can be compared against the three datasets MODIS, GlobCover and GLC-2000. MODIS describes this coordinate as urban area. As no annotation which falls into the urban category in Table 3 is given by any player, this dataset can be said to be false. GlobCover describes the given area as grassland, and GLC-2000 defines it as cultivated and managed. Although there are also 2 annotations given for grass, more annotations fall into the category of cultivated soil (fields, soil, land, agriculture, cultivated, farm, farmland, farmer). Therefore it can be said that the players agree most with GLC-2000 and least with MODIS.

The other 19 maps annotated by the players show very similar results. However, maps with a higher resolution sometimes deliver very special and detailed annotations. For example the Pyramids of Giza or the Colosseum have been recognized which makes it more difficult to use the annotations for improving land cover. In addition to this, the absolute number of occurrences does not always give information on the land cover distribution of a map. This means that not only the top but more annotations have to be taken into account for improving global land cover. The collected annotations have already been useful for evaluating the land-cover classes used in other games. As an example, for "stone," the players more often inserted "rocks", "mountain" or "hills," which seem to be more common names. For other categories, very specific annotations for well recognizable land cover like "river" or "lake" as well as "houses" and "streets" were given. This raises the question whether the other Landspotting games should use a broader set of land-cover classes. Interestingly, only for "cultivated soil" no other useful annotations have been collected.

5.2 Mean Error of the Generated Map

The Landspotting strategy game has already been played by 97 players, who have explored 7,756 km^2(1,939 tiles) land in total, of which 4,656 km^2 overlap. One tile is a high-resolution drawing (512×512 pixels) of 4 km^2 land (1 pixel represents 4 meters). The land-cover data painted by the user thus contains much more detailed information than simple annotations. In order to find out how much players agree on the explorations, we selected all coordinates (82) for which at least 3 drawings from different players exist. For each coordinate, we created an average map and calculated the mean absolute error, depicted in Figure 8. For at least half of these explorations, the average error lies under 10%, and for 64 tiles (76%), the error lies under 20%. This means that for half of all explorations, the players comply with over 90%. The small error of 10% occurs in most cases because the players explore the land in different qualities. Some players draw the map with very high detail, while others paint the map more quickly. For 76 coordinates, the error is still under 40%, which means that the players agree with 60%. This error is high because for some maps, it is quite difficult to distinguish between some of the surfaces, for example trees and bushes. But still, a lot of meaningful information can be extracted from this data. If, for example, no player has marked the land as "urban area" the land cover most likely does not contain an urban area. The step from 76 to 80 is quite noticeable, after which huge disagreements up to 100% exist. This occurs because some players have not taken the exploration task seriously and have just randomly painted the maps. This is possible because the output set (land-cover classes) is small and therefore players can still progress in the game even if they just randomly paint the earth, because even in that case, they sometimes agree with the other players and therefore still progress in the game. Therefore, as one gained insight, the game has to be much stricter, especially at the beginning of the game, to either get rid of such players as quickly as possible or by introducing a trust function for each player, which is used to decide whether their data should be ignored.

5.3 Comparison against Land-Cover Data Sets

For the following discussion, obviously wrong drawings have been removed from the dataset. Figure 7 shows the different land-cover maps of Japan (35.7556 latitude, 132.3125 longitude, 34.7683 latitude, 135.0869 longitude). The first map has been created from the strategy game. Note that only a fraction of the map has already been explored by the game, but the resolution of the map compared to the other land-cover maps is much higher, as one pixel of the map corresponds to 4 meters. The second map is MODIS, where one pixel corresponds to 500 meters. The third map is GlobCover, with a pixel size of 300 meters per pixel, and the fourth is GLC-2000, with a pixel size of 1,000 meters per pixel. As can be seen, the four land-cover maps highly disagree on some areas.

In order to gain more insight into which of the four land-cover maps performs better, they have been compared against 50 expert validation points, depicted on the last map in Figure 7, which lie on the land cover of the game map so that the validation points can always be compared against all 4 land cover maps. 48 of the 50 expert validation points agree with the land cover produced by the game. Only in two points the experts do not agree with the players. The experts mark the area as tree cover, whereas the players chose bushes, which seems to be very difficult to distinguish in these points. 42 of 50 validation points agree with MODIS. In some areas (1, 2, 3), MODIS does not correctly validate the land by either over- or underestimating urban or cultivated land, as has already been shown by the TAGinator game in Figure 6. In those regions, the game performs better than MODIS, also because of the much smaller pixel size used in the game. GLC-2000 agrees with 37 validation points, GlobCover with 36 validation points. Both products produce nearly the same result. In some areas (4, 5, 6), they both simply classify huge areas wrong by validating urban area as cultivated land. In one area (1), GLC-2000 performs a bit better than GlobCover, where GlobCover underestimates cultivated land. As this result shows, players are indeed able to improve global land-cover datasets, as the land-cover map created from the game receives better scores as the other land-cover products.

5.4 Player Reception

In this section, we present a small user study of how the main Landspotting game is currently being received by the players and what has to be further improved to attract a bigger audience. The user study consists of 34 random students who have been asked on Twitter to play the game for 2 weeks and to then give some feedback. As incentive to participate, the top 3 players received an USB flash drive.

In order to gain an objective result, some statistical information has been gathered. Figure 11 shows the gameplay time, Figure 9 how often the players logged into the game and Figure 10 depicts how much terrain has been explored by the players. One player started the game 5 times, 8 players at least 3 times, 15 players 2 times and the rest 1 time. In terms of explored terrain, one player explored over 400 km^2 terrain, the second-best player explored over 150 km^2 terrain, and the top ten players explored nearly 50 km^2. It is interesting that these numbers form a perfect logarithmic function. Not surprisingly, the time spent on the game, the number of logins and the land explored correlate. The longer the players play, the more land they explore and vice versa. Although we have tried hard to keep the players in the game as long as possible, especially at the beginning of the game by adding an advisor who introduces the player slowly to the game, the game is not yet able to keep a majority of players (50%) for more than 5 minutes in the game. From this, one can already conclude that the first impression of the game does not yet "hook" the majority of people and

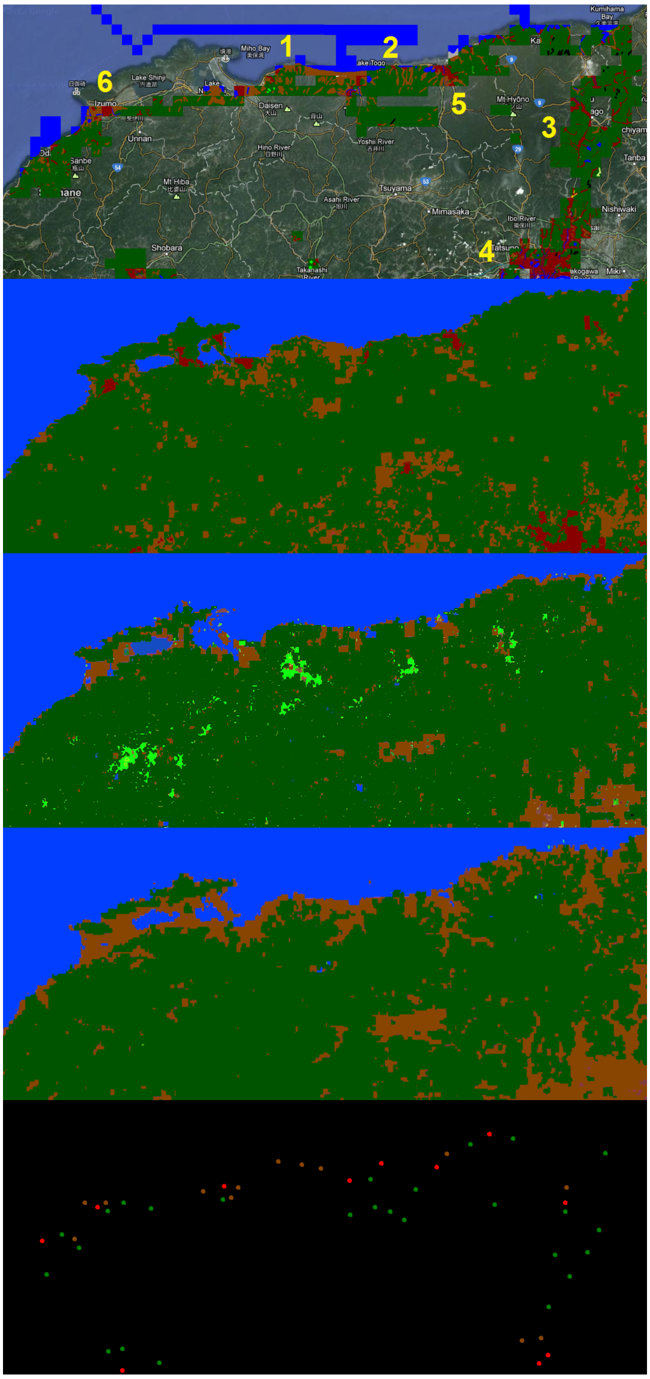

Figure 7: The topmost land cover is created from the game, the second is MODIS, then GlobCover and the fourth is GLC-2000. Dark green represents tree cover, bright green grass, brown cultivated and red urban area. The last map shows the 50 expert points against which the 4 land cover maps are compared.

that the game has to be further improved to reach the goal of creating a game for the masses. But still, this early version of the game has been able to entertain the top player for over 3 hours, the top three players for 1 hour, and the top 8 players for at least 20 minutes, while gathering valuable landcover data.

In general, all students seemed to like the game and game concept. The students stated that the graphics should be further improved and that more animations should be added to appeal to a bigger audience. The graphics are for some players also too cartoonish for a strategy game. The game is now too easy to "power-level" by just randomly exploring land, which we already discovered by analyzing the explorations. It was also mentioned that the scoring algorithm should be improved because as it is now, one possible cheating strategy is to just use the fill tool to fill out the whole area. Some game mechanics like harvesting have been received as confusing and should be removed from the game. Another important point which was made was that more tasks and midterm goals should be added to keep the players engaged for a longer amount of time, because a multiplayer game is boring at first until a critical mass of users has been reached. The idea came up to limit the area of where the players can explore the land so that the players are forced to play against each other in order to further increase the social factor.

Based on this user study it can be concluded that the following points have to be revised to appeal to a bigger audience:

- Graphics: done by a professional graphics designer.
- Gameplay changes: removing confusing game elements and adding some game mechanics to limit the gamefield to increase player to player interactions.
- More game tasks and midterm goals.

6. CONCLUSION

Four games for improving global land-cover have been presented. As has been shown, the players are able to create a land-cover map which receives better expert ratings than the three most-used land-cover products from NASA, ESA, and GEM today. Therefore, games can be a useful way to improve current land-cover datasets. In order to make the games fun and engaging, we have used a disjoint human computation game design by emulating already successful games like Civilization or Plants vs. Zombies and by adding the task Landspotting very harmonically to these classic games. In addition, we have made extensive use of social game mechanics. A small user study has revealed many important issues. For example, output agreement alone is not sufficient to avoid cheating strategies, as some people just randomly draw maps. Therefore, further control mechanisms like constantly checking against correct reference data are needed, especially at the beginning of the game. In addition, the graphics have to be improved further and more tasks have to be added to the game to reach a bigger audience and to keep the players engaged for a longer amount of time. We have not publicly announced the Landspotting games yet. We want to slowly increase the user base by iteratively improving the games based on the feedback of the users.

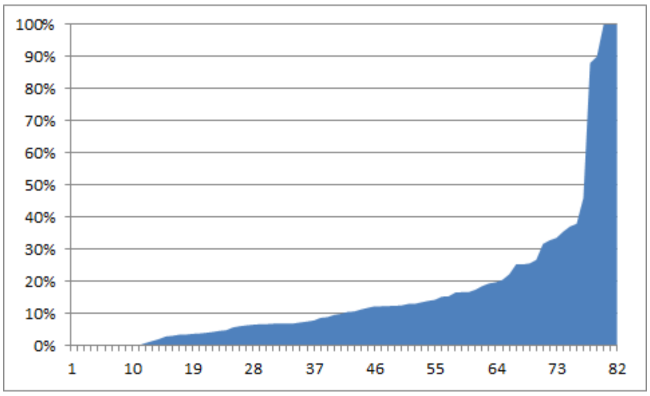

Figure 8: Mean absolute error of the 82 coordinates with at least 3 land-cover drawings.

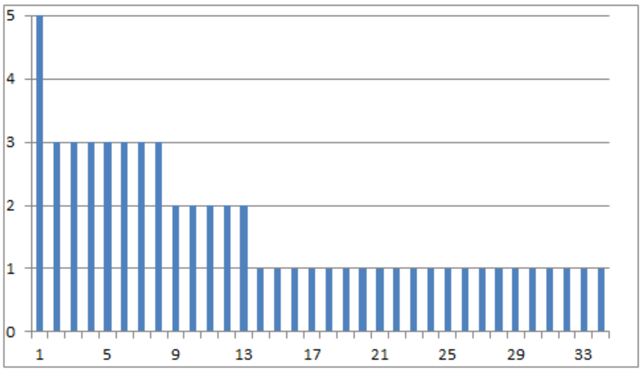

Figure 9: The number of logins by the 34 players.

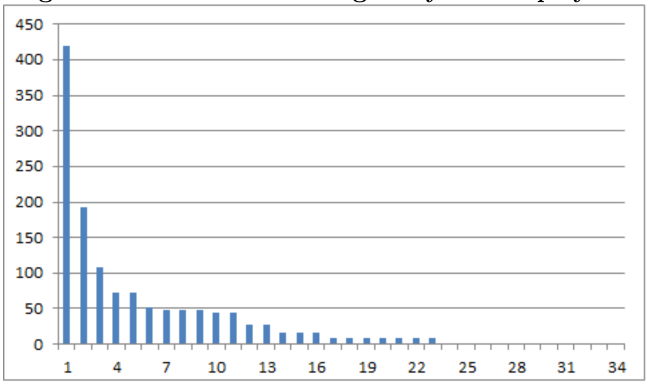

Figure 10: The area of land (in km^2) explored by the 34 players.

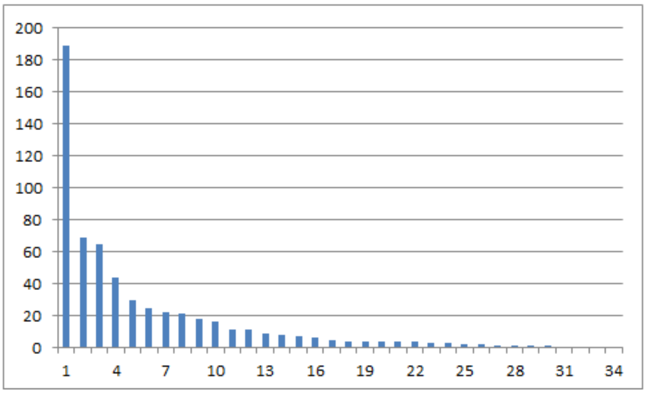

Figure 11: Durations (in minutes) the 34 players have played the Landspotting game.

7. ACKNOWLEDGMENTS

We would like to thank FFG for funding the Landspotting project. We would also like to thank our project partners IIASA, Fotec GmbH and EOX It Services GmbH for the great collaboration and, most importantly, we would like to acknowledge and thank all the Landspotting players!

8. REFERENCES

[1] L. Ahn and L. Dabbish. Labeling images with a computer game. *Proceedings of ACM SIGCHI*, pages 319–326, 2004.

[2] L. Ahn, M. Kedia, and M. Blum. Verbosity: A game for collecting common-sense knowledge. *Proceedings of ACM SIGCHI*, pages 75–78, 2007.

[3] G. Anthes. Html5 leads a web revolution. *Communications of the ACM*, 55:16–17, July 2012.

[4] Ho Chien-Ju, Chang Tao-Hsuan, Lee Jong-Chuan, et al. Kisskissban: A competitive human computation game for image annotation. *ACM SIGKDD Newsletter*, 12:21–24, June 2010.

[5] S. Cooper, F. Khatib, A. Treuille, Z. Popovic, et al. Predicting protein structures with a multiplayer online game. *Nature*, 466:756–760, 2010.

[6] S. Cooper, A. Treuille, J. Barbero, Z. Popovic, et al. The challenge of designing scientific discovery games. *FDG*, pages 40–47, 2010.

[7] S. Fritz, I. McCallum, C. Perger, et al. Geo-wiki.org: The use of crowdsourcing to improve global land cover. *Remote Sensing*, pages 345–353, August 2009.

[8] S. Fritz, L. See, I. McCallum, C. Schill, M. Obersteiner, M. Velde, et al. Highlighting continued uncertainty in global land cover maps. *Environmental Research Letters*, 6:4:044005, 2011.

[9] F. Khatib, S. Cooper, D. Baker, and others. Algorithm discovery by protein folding game players. *Proceedings of National Academy of Sciences USA*, 2011.

[10] M. Krause and J. Smeddinck. Human computation games: A survey. *Proceedings of the 19th European Signal Processing Conference*, September 2011.

[11] M. Krause, A. Takhtamysheva, M. Wittstock, and R. Malaka. Frontiers of a paradigm: exploring human computation with digital games. *Proceedings of the ACM SIGKDD Workshop on Human Computation*, pages 22–25, 2010.

[12] E. Law, L. Ahn, R. Dannenberg, and M. Crawford. Tagatune: A game for music and sound annotation. *Proceedings of ISMIR*, pages 361–364, 2007.

[13] C. Lintott, K. Schawinski, J. Vandenberg, et al. Galaxy zoo. *Monthly notices of the Royal Astronomical Society*, 389:1179–1189, 2008.

[14] J. Paavilainen. Critical review on video game evaluation heuristics. *Proceedings of Future Play*, pages 56–65, 2010.

[15] L. See and S. Fritz. A method to compare and improve land cover datasets. *IEEE Transactions on Geoscience and Remote Sensing*, pages 1740–1746, July 2006.