# Evaluating Risk to People and Property for Aircraft Emergency Landing Planning

Pedro F. A. Di Donato[*]

*National Civil Aviation Agency, 12246 São José dos Campos, Brazil*

and

Ella M. Atkins[†]

*University of Michigan, Ann Arbor, Michigan 48109*

**General aviation and small unmanned aircraft systems are less redundant, may be less thoroughly tested, and are flown at lower cruise altitudes than commercial aviation counterparts. These factors result in a higher probability of a forced or emergency landing scenario. Currently, general aviation relies on the pilot to select a landing site and plan a trajectory, even though workload in an emergency is typically high, and decisions must be made rapidly. Although sensors can provide local real-time information, awareness of more distant or occluded regions requires database and/or offboard data sources. This paper considers different data sources and how to process these data to inform an emergency landing planner regarding risks posed to property, people on the ground, and the aircraft itself. Detailed terrain data are used for selection of candidate emergency landing sites. Mobile phone activity is evaluated as a means of real-time occupancy estimation. Occupancy estimates are combined with population census data to estimate emergency landing risk to people on the ground. Openly available databases are identified and mined as part of an emergency landing planning case study.**

## Nomenclature

| | | |
|---|---|---|
| $c$ | = | country code |
| $D$ | = | day of the week |
| $d$ | = | date |
| $G_{ls}$ | = | set of candidate landing sites |
| $g$ | = | grid cell number |
| $h_{10\,m}$ | = | starting time of 10 min time interval of the day |
| $n_c$ | = | total number of country codes |
| $n_g$ | = | total number of grid cells |
| $n_t$ | = | total number of time intervals |
| $R_a$ | = | risk based on landing area |
| $R_h$ | = | risk to people on ground |
| $R_{ls}$ | = | overall landing site risk |
| $R_p$ | = | risk to property on ground |
| $R_v$ | = | risk to vehicle |
| $t$ | = | time, min |
| $t_{10\,m}$ | = | starting time of 10 min time interval, min |
| $W_h, W_p, W_v, W_a$ | = | weights associated with landing site risks |
| $w$ | = | $A^*$ transition cost weight |
| $\Theta$ | = | median of maximum aggregated SMS and call activity during a day, for a day of the week |
| $\lambda_{census}$ | = | occupancy from census data |
| $\hat{\lambda}_{census}$ | = | normalized occupancy from census data |
| $\lambda_{rt}$ | = | real-time occupancy estimation |
| $\lambda_0$ | = | $A^*$ fixed transition cost element |
| $\Phi$ | = | median of aggregated SMS and call activity for each grid cell, day of the week and time interval of the day |
| $\Phi_{grid}$ | = | sum of $\Phi$ for all grid cells |
| $\hat{\Phi}$ | = | normalized $\Phi$ |
| $\phi$ | = | aggregated SMS and call activity |
| $\hat{\phi}$ | = | normalized $\phi$ |
| $\phi_1$ | = | activity in terms of SMS-in |
| $\phi_2$ | = | activity in terms of SMS-out |
| $\phi_3$ | = | activity in terms of call-in |
| $\phi_4$ | = | activity in terms of call-out |
| $\phi_5$ | = | activity in terms of internet |
| $\phi_i$ | = | modified activity feature for heat map plot |

$\phi_{i_{\mathrm{avg}}}$        =    average activity in terms of feature $i$ throughout the entire database

$\phi_{\mathrm{grid}}$        =    sum of $\phi$ for all grid cells

## I. Introduction

**T**RIPLY redundant commercial transport aircraft are rarely forced to land off-runway, but emergency landings are occasionally required today in general aviation and experimental aircraft operations. General aviation (GA) and small unmanned aircraft systems (UAS) cannot rely on triply redundant systems due to cost and weight considerations. Furthermore, GA and UAS operations will not be cost-effective should a team of two well-trained pilots be required to continuously supervise each flight, reducing pilot decision redundancy relative to commercial transport operations. UAS will likely require autonomous safe emergency landing capability before large-scale integration into the National Airspace System, particularly when operating over populated areas. GA and UAS applications, therefore, motivate further research in sensing, data mining, and autonomy algorithms that reduce risks for scenarios requiring a forced or emergency landing.

Emergency landing is an important case in which aviation is forced to interact with the "nonaviation" world. If landing on a conventional runway or helipad is not a feasible option, pilots must now choose a landing site based on what they can spot through the aircraft windows, prior knowledge of the nearby terrain, and (at best) a hurried examination of charts. Current cockpit avionics may display the closest airport runways as well as other visual navigation landmarks (e.g., radio towers and buildings) but do not propose off-airport landing sites. Emerging information sources offer new opportunities to improve flight safety. For example, automatic dependent surveillance–broadcast (ADS-B) receivers are currently available to provide local air traffic data in support of detect-and-avoid automation. Although ADS-B is a significant advance, aviation need not remain insular with respect to data feeds. Current data storage and transmission technology opens a whole new world of information sources that can be exploited in real time to improve aviation safety and to address new challenges unmanned aviation will introduce [1].

Emergency landing should take into account the fact that the risk to people on the ground is a function of dynamic day-to-day events as well as phenomena captured in a static database. For example, an empty soccer field can be an excellent small UAS emergency landing area. However, a school soccer field is likely to be populated during school hours and during an after-school practice or game. Time of day generally impacts risk as a function of overflown region type. For example, districts with a high concentration of bars and nightclubs are expected to have high occupancy on a Friday night but relatively low occupancy on a Monday morning.

Different data sources can be combined to maximally inform decision algorithms for emergency landing. Onboard sensors can provide real-time information [2]. Such information, however, is restricted by sensor field of view, occlusion, and range constraints. Onboard databases can provide an emergency landing planner information over a much larger reachable landing area [3]. Data link enables updated information to be accessed in real time. Databases and data link can also be used to determine low-risk UAS flight paths or for UAS highway definition as proposed by [4].

This paper proposes the Fig. 1 algorithm for exploiting onboard sensors, static database information, and real-time data link for emergency landing. This approach augments the decision strategy and static data proposed in [3] with data link and a list of specific landing sites vetted during preprocessing. In the case of an emergency, the onboard decision algorithm can fuse database and data-link information with locally acquired sensor data to decide between landing locally based on onboard sensor feedback versus flying to a known landing site beyond immediate sensor field of view but inside the reachable area or footprint, considering the aircraft's remaining range estimate. Candidate landing sites can be stored before flight in an onboard database with an associated risk model that may be augmented in real time with data-link information. Planning the path to the lowest-risk landing site must also take into account and minimize risk exposure during the flight using overflown region information from onboard databases with possible real-time data-link augmentation. Once the aircraft reaches the vicinity of the candidate landing site, onboard sensors can again be used to assess local risk while an alternate landing site may also be considered for cases where the originally planned site would introduce unexpectedly high risk. The limited amount of stored energy must be taken into consideration to ensure that the aircraft eventually chooses to land rather than crashing en route once energy stores are fully depleted.

This paper investigates nontraditional data sources for emergency landing site selection and planning. In a previous publication, the authors explored this idea through the incorporation of a road database to identify potential off-field emergency landing sites [5]. This work more generally examines data sources that can be used to define candidate landing sites and improve risk estimates through data-link information transmission. This work complements robotics and aerospace publications focused on real-time sensor data processing for local risk estimation and planning; both are essential to safe autonomous emergency landing. Contributions of this work are focused on the two dashed bordered
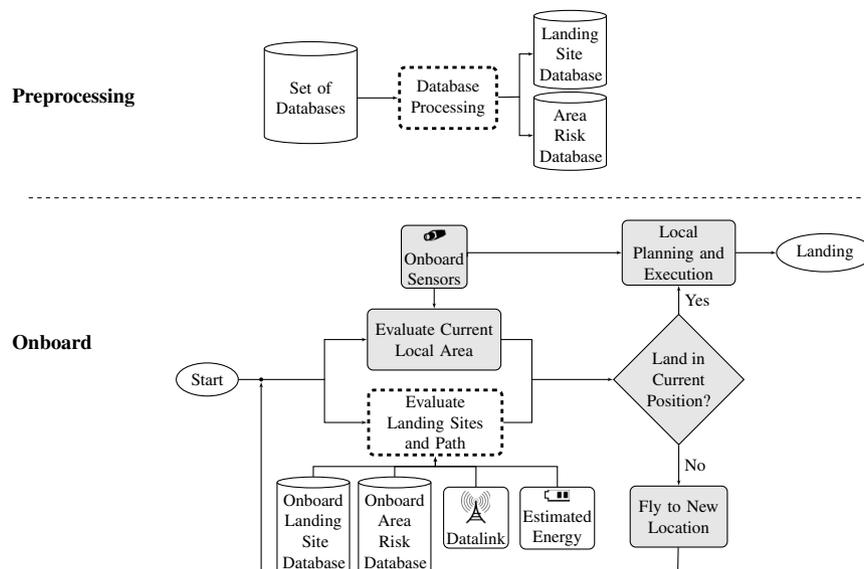


**Fig. 1  Proposed decision logic for autonomous aircraft emergency landing. This paper exploits novel data link and database resources to define minimum-risk landing sites and approach paths.**

**Table 1    Nonaviation database sources used in previous work**

| Information type | Reference | Source |
|---|---|---|
| Digital elevation map | [2,11] | Shuttle Radar Topographic Mission |
| | [3] | U.S. Geological Survey, National Elevation Set |
| Man-made structures/buildings | [3] | New York City Primary Land Use Tax-Lot Output |
| | [10] | Ordinance Survey of Northern Ireland |
| | [11] | OpenStreetMap |
| Surface/terrain type | [3] | U.S. Geological Survey, National Land Cover Dataset |
| | [10] | Ordinance Survey of Northern Ireland |
| | [11] | OpenStreetMap |
| Roads | [5] | OpenStreetMap |
| Population | [3] | U.S. Census |

modules of Fig. 1. The first contribution of the paper is to explore an emerging information source, the mobile phone activity database, for occupancy estimation. This idea has received recent attention outside aerospace [6–9] but to date has not been applied to aviation.

The second contribution of the paper is to propose a method to explore geographic information system (GIS) databases to identify candidate emergency landing sites and to evaluate their risks to people on the ground, property damage, and aircraft damage or loss. This paper takes full advantage of open data to provide contextual information that provides a more informed risk assessment of candidate landing sites than in previous work [1,3,10]. The proposed method allows consideration of temporal (time-of-day) influence on the computed risk of each landing site, something recognized but not used in [3], and that has been addressed at only a basic level in [1]. Moreover, this paper proposes the use of mobile phone data to estimate risk in the case of special events that increase the number of people in the vicinity of candidate emergency landing sites.

Third, this paper contributes the first fusion of mobile phone activity data with census data to provide a real-time estimate of risk posed to an overflown population. Such risk can then be stored in a table or matrix format potentially updated before each flight for use by an onboard flight planner [3,11]. Benefits and limitations of the examined databases and proposed risk evaluation strategy are analyzed and illustrated with a case study over the Milan, Italy, metropolitan area, for which a mobile phone dataset is openly available.

This paper is organized as follows. Section II presents previous work in emergency flight planning and use of mobile phone activity for occupancy estimation. Section III presents and analyzes the Milan, Italy, mobile phone activity database, whereas Sec. IV describes the population census and map databases used in this work. Section V presents the methods applied for emergency landing site identification and real-time occupancy estimate as well as flight path planning. Section VI presents a complete case study from data preprocessing steps through final emergency flight planning. Sections VII and VIII offer a discussion and conclusion, respectively.

## II.    Related Work

### A.    Emergency Flight Planning

The keys to successful emergency management are "aviate, navigate, communicate". Aviate translates to maintaining stable control of the aircraft. Navigate translates to assessing the situation with respect to the current flight plan. If the situation indicates a near-term emergency landing, a landing site must be selected, and a feasible flight plan to that site must be rapidly computed and executed. With data link, any aircraft declaring an emergency will have top priority supporting an assumption that deconfliction with other air traffic is not required.

For fixed-wing airplanes, emergency landings on airport runways are the first choice if they are reachable based on aircraft location and the nature of the emergency situation. Runway choice is usually performed using traditional aviation databases providing information such as runway properties, airport facilities, and wind [12]. The risk associated with the planned flight to each landing site can also be incorporated into the landing site decision [13].

If safe flight to a traditional runway is infeasible, an off-airport emergency landing site must be selected. One option proposed in particular for rotorcraft is to identify and select a landing site using only onboard sensors. Vision systems can be used to simply identify flat surfaces that are considered safe [14] or to make a more complex safe landing site decision based on a combination of texture and flatness [15]. Other researchers proposed edge detection as a basis for candidate landing site identification [2]. Landing site identification from a point-cloud model that can be generated from lidar sensors was also proposed [16].

The major disadvantage of using only sensors for landing site identification is their limited range and dependency on aircraft orientation [3]. The safest landing site can indeed be behind the aircraft, just beyond sensor range, or obscured by an obstacle. Although landing sites might be identified over the course of several flights based on sensor data, several available nonaviation or "nontraditional" databases can also assist in landing site selection. Table 1 presents an overview of nontraditional databases referenced in the aviation literature.

Several applications have been proposed for these databases. For example, Rackliffe et al. [1] combined different information sources in the form of a GIS data to produce a raster (i.e., a cost matrix) that can be used for landing site identification, selection, and flight planning. Such a matrix can be two-dimensional [11] or three-dimensional [3]. The case study presented in [5] relied only on databases to identify candidate landing sites directly from the raw vector data and used some of the properties of road segments identified as landing site candidates for utility evaluation. The algorithm proposed in [17] uses as landing site the point in the aircraft reachable area with minimum total landing cost based on different data such as population and land use information.

Another approach is to use databases as information sources for features not available from onboard sensors [2]. An integrated approach was proposed in [10], where features are found using a combination of sensor and database information. In their work, databases were used as prior knowledge in a maximum-likelihood classifier for terrain type.

During an emergency scenario, three risks are identified: endangering humans, property damage, and aircraft damage. For UAS, risk to humans only occurs when the UAS collides with a manned aircraft or crashes into a populated area. Previous work [10,11] has aimed to minimize risk at a potential emergency landing area. Table 2 presents a summary of factors identified in related research.[‡]

---

[‡]Mejias et al. [2] did not specifically define how each feature is associated with each risk factor. Table 2 was constructed comparing these reference data sources with other references that associate such data sources with a specific risk.

Table 2    Risk estimation for emergency landing scenario used in previous work

| Risk | Reference | Assessed using |
|------|-----------|----------------|
| People on ground | [2,10] | Distance to man-made objects |
|  | [3] | Population density |
|  | [11] | Distance to buildings |
| Property damage | [2,10] | Distance to man-made objects |
|  | [3] | Distance to buildings |
|  | [11] | Distance from power lines and transportation ways |
| Aircraft damage | [2,3] | Terrain slope and type |
|  | [10] | Terrain type and roughness |
|  | [11] | Terrain slope and roughness |

Two observations are made regarding previous work to estimate risk to people on the ground. First, multiple researchers have relied on the distance to man-made objects or buildings, a collision metric. A collision metric can translate to UAS accident risk but does not carry information about the number of people potentially exposed to a UAS accident. Second, most risk metrics in previous work are static (constant) over time, whereas the human occupation of an area can vary substantially over time. Rackliffe et al. [1] described how GIS information could be used to provide a time-dependent risk estimate based on a binary safe/not-safe raster classification. For example, sports arenas might be considered unsafe during weekend events and safe during the regular week.

Once landing site candidates are identified, a flight plan to each site can be constructed. Landing sites can be prioritized before flight planning [5,12] with subsequent focus on a single site, or sites can be prioritized with additional consideration of flight plan risks after the plan is computed [3,11,13]. Flight planning can be based on database, sensor, or combined information. As in landing site selection, there is a tradeoff between reliance on real-time information within the local field of view versus consideration of longer-range information, which can be outdated or incomplete.

### B.  Mobile Phone Databases for Population Estimation

Mobile phones are one of the most successful innovations of the new millennium. Their widespread usage and persistent network connectivity provide a new way to monitor real-time area occupancy with minimal overhead. Studies that rely on mobile phone traffic datasets appeared in 2006, and their use is increasing at a rapid rate. An extensive survey of this new research area is presented in [6]. The authors organize the studies in three main areas: social analysis, mobility analysis, and network analysis. Each has distinct subdivisions, ranging from epidemics to networking solutions.

There are different strategies for data collection from a mobile network [6]. One of the simplest and most explored datasets generated from mobile phones is call detail reports (CDRs). Mobile carriers generate one report each time a mobile phone makes a call or a transaction initially for billing purposes. Those reports consist of the caller ID, recipient ID, call time, and duration. They also include information regarding the cell tower used for the caller/recipient. This information can be processed to estimate the geographic position of the phone during the call. When multiple towers are detectable from a mobile phone, positioning techniques such as observed time difference of arrival can also be used to improve cell tower geolocation [18].

Once user locations are determined, it is possible to estimate the population density or occupancy based on aggregated mobile traffic. Researchers have established connections between the underlying population distribution obtained from census estimates with telecommunications data [19]. A recent study demonstrated that "presence" information (i.e., an estimate of the number of users in a certain space based on their last interaction with the mobile network) can provide better correlation with census data [7]. The same study also proposed a method to evaluate real-time population distribution and showed reasonable results, although ground truth was not available. One method to independently evaluate real-time occupancy or presence was proposed using soccer match attendance and the number of flights arriving and departing [8]. Soccer stadium results were reported better than airport results, although the authors did not determine whether the model developed for one particular space such as the stadium is valid for other areas of the city. A more detailed analysis of the effect of a soccer game on the mobile network was performed in [20]. Detecting events from unusual mobile phone activity was studied in [9]. The authors also evaluated the use of Twitter feeds for event detection and identification. Bagrow et al. [21] offer a detailed study of the impact of events on network usage including differences between planned and unexpected events.

Real-time occupancy is also related to human mobility. The survey of Naboulsi et al. [6] also includes work in this area. Isaacman et al. [22] proposed a model of human mobility in a metropolitan region based on CDR and propose and evaluate the use of census data when a CDR database is not available. Although work to date has applied mobile traffic data to mobility studies [6], these results also support the use of mobile phone activity data to estimate risk to the local population during an emergency aircraft landing. This paper proposes the fusion of mobile phone activity with other data sources for overall risk assessment within an emergency landing flight planner.

## III.   Mobile Phone Activity Database

Although companies worldwide are collecting and processing large-scale mobile phone data for internal purposes, many of these databases are not openly available due to data privacy and proprietary concerns. Cellular phone service providers are particularly sensitive to data sharing for these reasons. Fortunately, companies sporadically open data sets to support research and education challenges (e.g., Orange's Data for Development and the Telecom Italia Mobile Big Data Challenge). The latter competition's 2014 data became an open database in December 2014,§ providing a rich source of data analyzed to provide a mobile phone occupancy estimation metric for this paper.

### A.   Telecom Italia Mobile Big Data Challenge 2014 Database

The Telecom Italia Mobile (TIM) open database contains information for two Italian areas: the Milan metropolitan area and the Trentino territory. In addition to cellular usage data, it contains time-stamped data such as weather, precipitation, air quality, social media data, news, and, in the case of Trentino territory, electrical current usage. This initial investigation of mobile phone data usage for aviation only considers telecommunication activity datasets for the city of Milan collected over the months of November and December 2013.

---

§"Dandelion API — Open Big Data," https://dandelion.eu/datamine/open-big-data [retrieved 12 December 2015].

The TIM database is generated by aggregating data from CDRs. Spatial and temporal aggregation combined with unit-of-measure blurring are used for privacy and proprietary reasons [19]. Database entries are summarized next (see footnote §), including definitions of associated symbols.

1) *Location*: $g$, grid cell number. Data are provided for each of 10,000 square cells covering the Milan region. Each cell has size $235 \times 235$ m. Grid geographical locations are available in a separate database defining each square grid element by its four vertex or corner coordinates.

2) *Time interval*: $t_{10\,m}$, a time stamp in UNIX time marking the beginning of a 10 min period over which the reported data were collected. This work assumes that $t_{10\,m}$ is already preprocessed to represent local Milan time in minutes, where $t_{10\,m} = 0$ represents an interval starting at midnight (0000 hrs), 1 November 2013.

3) *Country code*: $c$, country of origin for the phone receiving/sending the communication.

4) *SMS-in*: $\phi_1$, activity in terms of the received short message service (SMS) inside grid tile $g$ during time interval $t_{10\,m}$; data are sent from the nation identified by country code $c$.

5) *SMS-out*: $\phi_2$, activity in terms of transmitted SMS, inside grid tile $g$, during time interval $t_{10\,m}$ and received by the nation identified by country code $c$.

6) *Call-in*: $\phi_3$, activity in terms of received calls inside grid tile $g$, during time interval $t_{10\,m}$, and issued from the nation identified by country code $c$.

7) *Call-out*: $\phi_4$, activity in terms of issued calls inside grid tile $g$, during time interval $t_{10\,m}$, and received by the nation identified by country code $c$.

8) *Internet*: $\phi_5$, activity in terms of Internet traffic inside grid tile $g$, during time interval $t_{10\,m}$, and by the nation of the users performing the connection identified by country code $c$.

The original CDRs used to compute $\phi_i$ are generated every time a phone receives or sends an SMS ($\phi_1$ and $\phi_2$), every time a phone receives or issues a call ($\phi_3$ and $\phi_4$), and every time a user starts or stops an Internet connection plus every time 15 min or 5 MB of usage occurs since the last CDR ($\phi_5$) (see footnote §).

### B. Mobile Phone Activity Data Analysis

After decompression, the first database processing step was to aggregate datasets over all country codes because CDR-based grid occupancy estimates should not be biased by country of information origin, user origin, or destination. Aggregation represents the following computation:

$$\phi_i(g, t_{10\,m}) = \sum_{c=1}^{n_c} \phi_i(g, t_{10\,m}, c) \tag{1}$$

where $n_c$ is the total number of country codes in the database. To reduce data storage and transmission overhead, the database omits $\phi_i(g, t_{10\,m}, c)$ for each case where no data exchange was performed during one of the considered time periods. This may result in $(g, t_{10\,m})$, for which $\phi_i(g, t_{10\,m}, c) = 0, \forall\, i \in \{1, 2, 3, 4, 5\}$, being omitted from the data. The aggregation step inserts zero-value $(g, t_{10\,m})$ lines as needed to fully populate the data set. This processing step aggregates mobile activity data into a single data file with 1,440,000 lines of seven entries for each 24 h day.

Because numerical entries in the open database are blurred to address privacy and proprietary data concerns,[¶] the only information available is that the provided data respect trends (e.g., a higher SMS-in number indicates a higher number of received SMS). Moreover, although the scale of SMS and calls are comparable, Internet access scales differently (see footnote §). To plot grid occupancy estimates from telecommunication exchanges as a heat map, the maximum $\phi_{i_{\max}}$, minimum $\phi_{i_{\min}}$, and average $\phi_{i_{\text{avg}}}$ value for each grid or field across the entire database were computed and recorded. For example, $\phi_{i_{\text{avg}}}$ is computed from

$$\phi_{i_{\text{avg}}} = \frac{\sum_{g=1}^{n_g} \sum_{t_{10\,m}=1}^{n_t} \phi_i(g, t_{10\,m})}{n_g n_t} \tag{2}$$

where $n_g$ and $n_t$ are the total number of grid cells and time intervals, respectively. During processing, it was observed that the $\phi_i$ intervals are too large to support linear interpolation between the minimum and maximum value for each cell. Instead, a value of 40 times the overall average of the field was set as a maximum or saturation limit for both interpolation and clipping of outlier values. Mathematically, heat maps are created with the following modified features $\bar{\phi}_i$:

$$\bar{\phi}_i(g, t_{10\,m}) = \min\left(\frac{\phi_i(g, t_{10\,m})}{40\phi_{i_{\text{avg}}}}, 1\right) \tag{3}$$

This value was considered adequate to generate heat maps for each $(g, t_{10\,m})$ pair. Figure 2 presents a typical heat map constructed with database feature call-out ($\phi_4(g, t_{10\,m})$) with 10 min local time interval starting at 10:00 a.m. presented in the figure as an image overlaid with Google Maps. This time shows intense activity commensurate with weekday business hours. The greatest intensities highlight downtown, Central Rail Station, and Milan Linate Airport terminal. Note that the available database region is larger than that presented in Fig. 2, but in this section, results are restricted to the depicted window to take full advantage of Google Maps labels.

The differences between Fig. 2 and both graphs shown in Fig. 3 reflect the differences between weekday, weekend, and night time periods on the volume of mobile communication activities. This observation reveals the first limitation of using mobile phone data for occupancy measurement. During the overnight hours, people reduce their phone use, but a significant fraction of these people remain in the city even if they "disappear" with respect to the CDRs. Figure 4 shows the sum of SMS and call total grid activity at each time interval of the database. The variation between weekdays and weekends is clear as well as variations within each day. Note that, for Milan, 1st November is a holiday and that Christmas week has a significantly different pattern.

Even though activity levels diminish, night-time mobile activity combined with land use data has been used to estimate occupancy or population using census data as a ground-truth comparison [19]. Khodabandelou et al. [7] proposed the use of mobile activity and census data to provide a real-time population estimate. A model was established using overnight values and census data. Then, such a model was also applied to

---

[¶]An aviation data provider would be better able to negotiate with cellular service providers to ensure that accurate (unblurred) data are represented while respecting privacy and proprietary concerns.
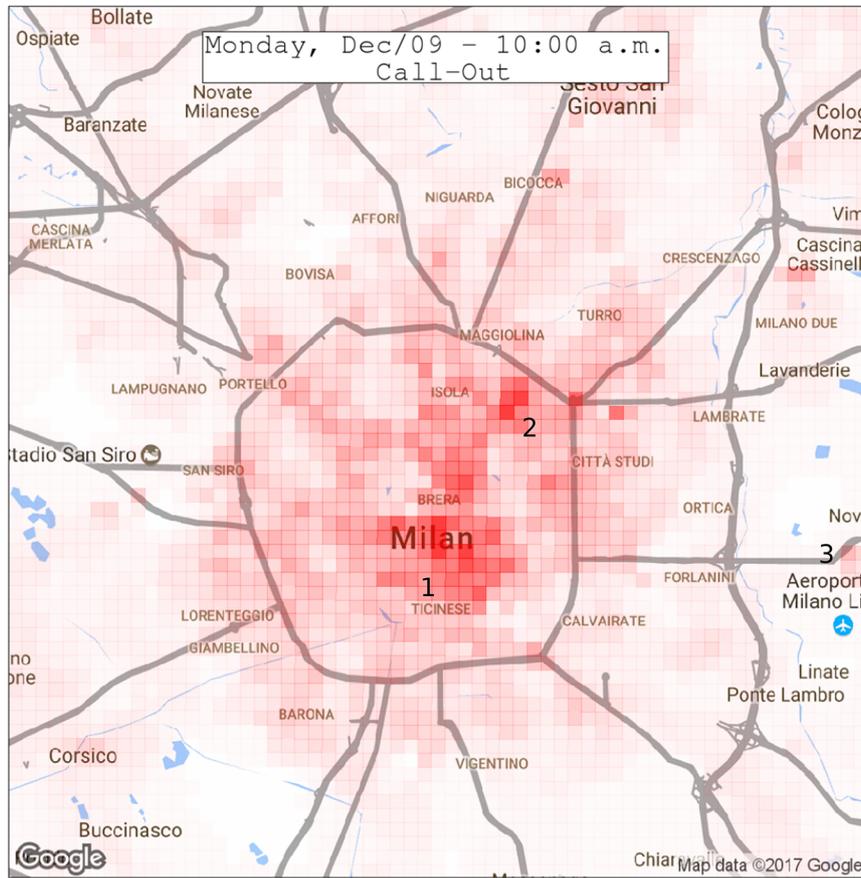
**Fig. 2    Typical business-hour mobile phone activity: downtown (1), rail station (2), and airport terminal (3).**
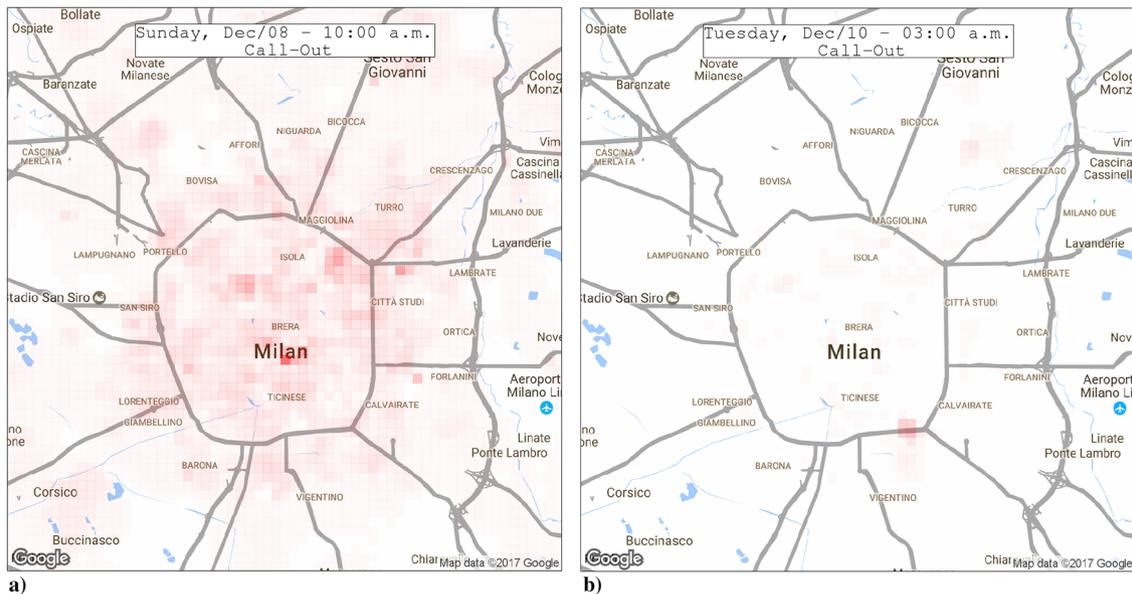
**Fig. 3    Low mobile phone activities related to weekend or late-night time intervals.**

other times of the day. Instead of switching between exclusive use of mobile phone activity or census data, our work proposes to fuse both datasets to estimate grid occupancy.

Figure 5 illustrates one of the main benefits of the mobile activity database: its ability to estimate real-time grid occupancies that can significantly differ from statistical (expected) values. On 22nd December 2013, a famous rivalry soccer match, Internazionale versus Milan, took place. Network activity on this date is presented in Fig. 5. Regions of high activity were noted just before, during the half-time interval (not shown), and just after the game. During the game, however, the region does not light up significantly even though a huge concentration of people occupies the stadium. Figure 6 presents the sum of SMS and call activities for the grid cell that represents the stadium plotted with respect to the time of the day for weekdays. Each line corresponds to one day. This figure shows the pattern for SMS and call activities in different game days and in contrast to other days. These results confirm the soccer match behavior observed in [20] for two games in Brazil as well as the gradual increase in activity before the game, a characteristic of planned events as per [21].
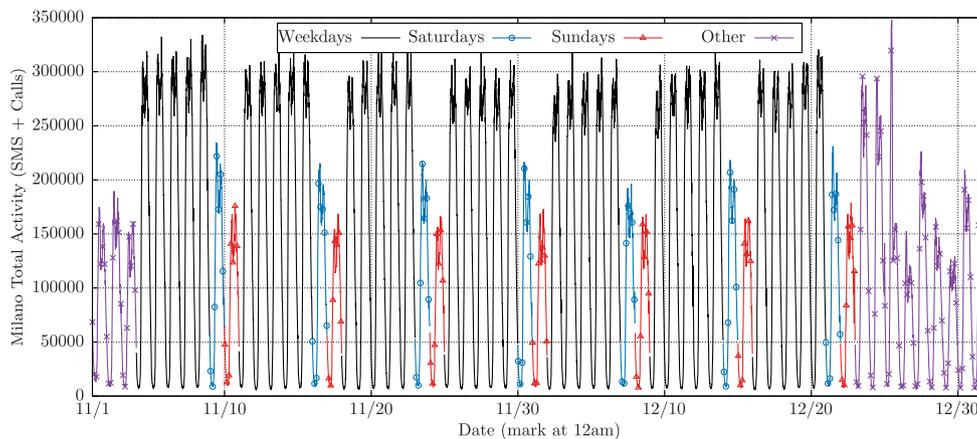
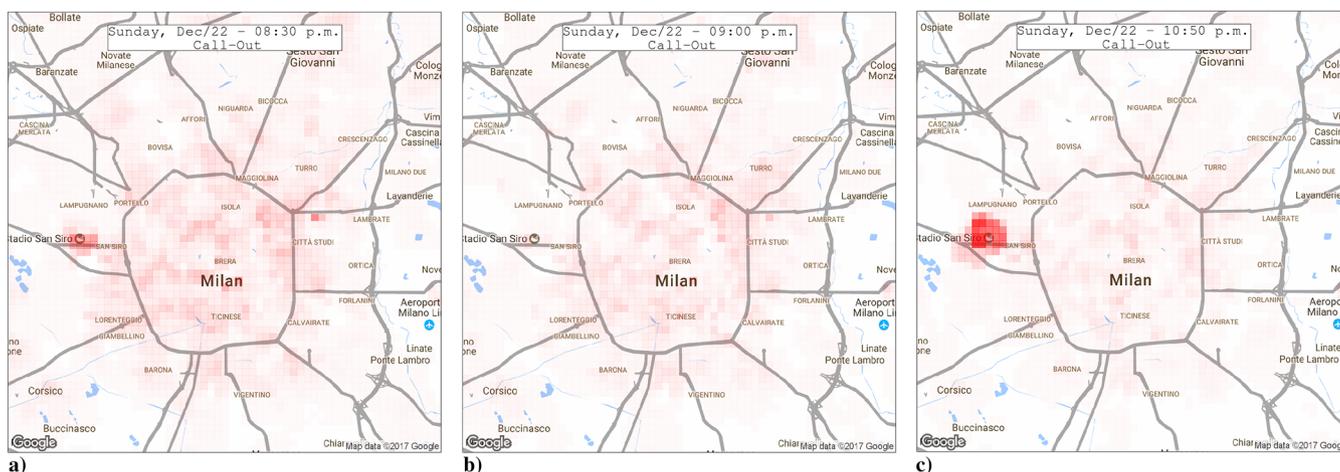**Fig. 4    Total mobile phone SMS and Calls activity throughout the database time period.**



**Fig. 5    Call-out activities during a soccer match: a) before start, b) during, and c) after end.**
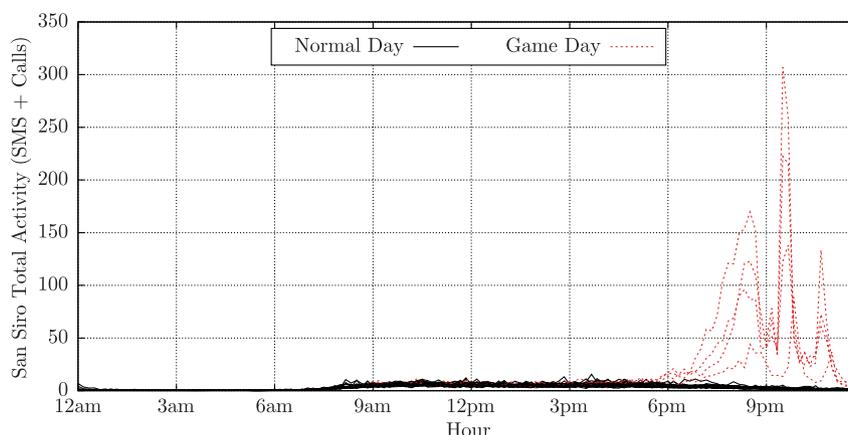


**Fig. 6    Total mobile phone activity at San Siro cell on different weekdays.**

High usage profiles would allow an aircraft planning an emergency landing to recognize and avoid this populated area despite a soccer field normally holding promise as a relatively safe off-runway landing site.[**] The use of a "presence" term discussed in Sec. II.B could mitigate the risk of low mobile phone activity periods being recognized as a low occupancy. Such data, however, are available only in the 2015 dataset, which is not yet publically available.

The heat maps presented previously used call-out feature $\phi_4$. This option follows [19], which reports that the call-out feature gave the highest correlation with population data from traditional census databases. Figure 7 shows heat map trends for other mobile activity features over the same after-match time interval as call-out in Fig. 5. Note that SMS-in and SMS-out clearly also show higher activity around the stadium areas. Call-in data are similar to call-out data and are thus not shown here. Internet activity, however, shows a significantly different map. In fact, Internet data is not considered a good proxy for human activity due to Internet traffic generated by applications running in the background [7].

---

[**]San Siro Stadium is partially covered, which poses additional constraints to an emergency landing, but it is used as a general example.
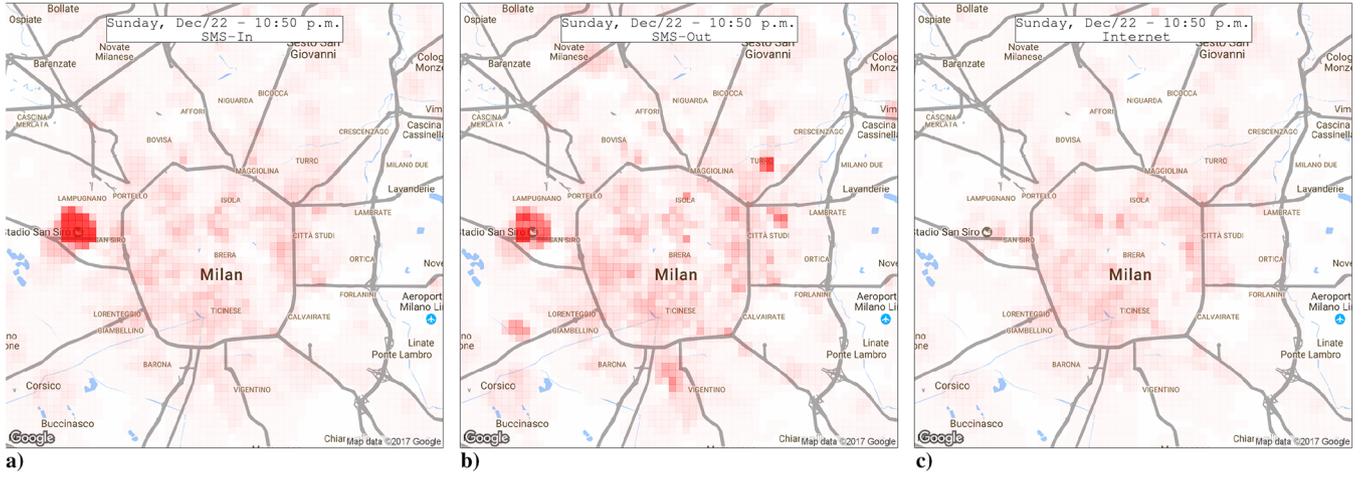
**Fig. 7 Activity in terms of different features just after a soccer match.**

## C.  Proposed Mobile Phone Activity Data Preprocessing

Although the use of "heat maps" gives good insight about the mobile phone activity database, a different preprocessing algorithm is proposed for application to landing site risk evaluation and real-time estimation of overflown area occupancy for the flight planner.

First, mobile calls and SMS data are aggregated per the original 10 min report intervals, an approach also used by other researchers [9,23]:

$$\phi(g, t_{10\,m}) = \sum_{i=1}^{4} \phi_i(g, t_{10\,m}) \tag{4}$$

Next, this work performs different time aggregations using three time-related variables. First, the time interval of the day $h_{10\,m}$ is defined as

$$h_{10\,m} = t_{10\,m} \bmod 1440 \tag{5}$$

Note that there are 144 ten minute intervals in a day. Second, date $d$ is defined for each day, which covers 144 particular values of $t_{10\,m}$. Third, days of the week $D \in \{D_{weekday}, D_{saturday}, D_{sunday}\}$ are defined as per Fig. 4. Each $D$ can be interpreted as a set of particular values of $d$. The "other days" values are not used because they would correspond to holidays, during which mobile phone activity is anomalous (see Fig. 4) relative to the other three groups.

The median of aggregated SMS and call activity is computed for each cell, each day of the week $D$, and each interval of the day $h_{10\,m}$ using an approach similar to the method from [9]:

$$\Phi(g, D, h_{10\,m}) = \underset{t_{10\,m} \in D, t_{10\,m} \bmod 1440 = h_{10\,m}}{\mathrm{median}} [\phi(g, t_{10\,m})] \tag{6}$$

Two other parameters are computed for real-time occupancy estimation. First, the total activity on the grid is computed as illustrated in Fig. 4:

$$\phi_{grid}(t_{10\,m}) = \sum_{g=1}^{n_g} \phi(g, t_{10\,m}) \tag{7}$$

Note that total activity can also be computed from median values:

$$\Phi_{grid}(D, h_{10\,m}) = \sum_{g=1}^{n_g} \Phi(g, h_{10\,m}, D) \tag{8}$$

The total activity figure is used to normalize the activities in each cell at a given time $t$:

$$\hat{\phi}(g, t_{10\,m}) = \frac{\phi(g, t_{10\,m})}{\phi_{grid}(t_{10\,m})} \tag{9}$$

The median values for each day of the week and hour of the day can also be normalized:

$$\hat{\Phi}(g, D, h_{10\,m}) = \frac{\Phi(g, D, h_{10\,m})}{\Phi_{grid}(D, h_{10\,m})} \tag{10}$$

Second, the median of the maximum total activity on each day for every day of the week is computed:

$$\Theta(D) = \underset{t_{10\,m} \in D}{\mathrm{median}} \left[ \max_{t_{10\,m} \in d} (\phi_{grid}(t)) \right] \tag{11}$$
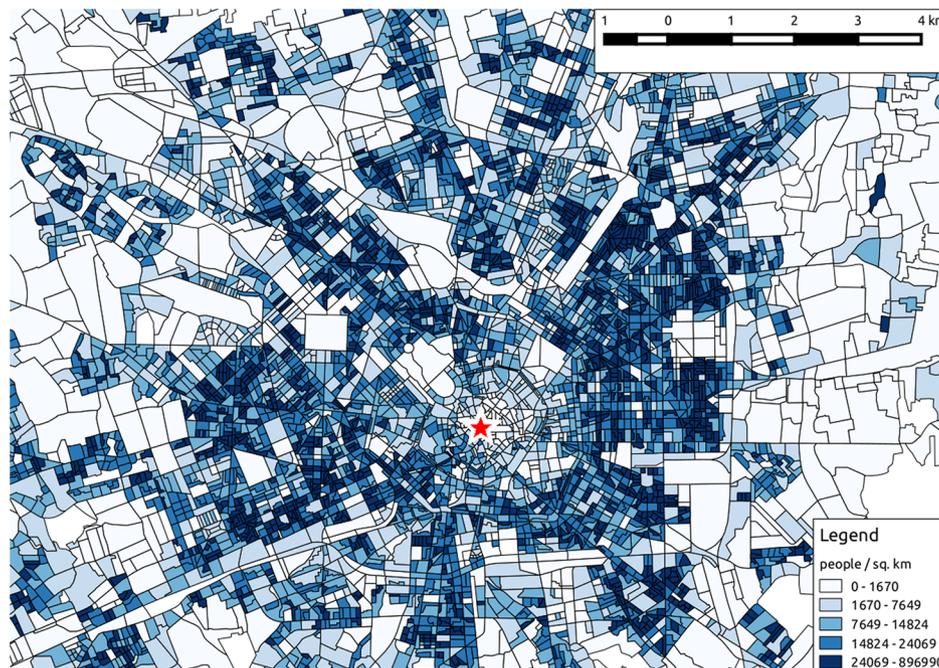
**Fig. 8   Population density for the central Milan region. The star marks downtown.**

This work uses median values instead of simple means. The median is also a measure of central tendency, but it is very insensitive to the presence of outliers [24]. Examples of data outliers are the high values during specific events such as soccer games illustrated in Fig. 6.

## IV.   Other Databases

This work fuses data from mobile phone activity with census data to estimate occupancy. Although similar databases have been used for nonaviation applications in previous work (see Sec. II.B), this paper proposes a novel fusion strategy as well as a new use case: estimation of the risk posed to people and property on the ground during a manned or unmanned aircraft emergency landing scenario.

### A.   Italy Census Data

The Italy 2011 Population and Housing Census ("Censimento della popolazione e delle abitazioni") is publicly available[††,‡‡] with data aggregated into different geographic categories. This work uses the most detailed divisions, "sezione di censimento", which approximately follow city blocks. This results in a nonuniform grid representation of the data. This work uses one of the available census parameters: "resident population — total".

Data are available for each administrative region of Italy. Preprocessing begins by downloading census table results and census geographic divisions. This paper considers census data within a boundary around the Milan area where mobile phone data are also available. The census grid is clipped to the zone of interest, and data from census results are incorporated into the appropriate census data cells. For illustration, population density was computed from each census grid's population divided by the grid's area. Results are presented in Fig. 8, where downtown Milan is marked by a star. Note that there are different areas with low population density close to downtown. Although some areas represent parks and open areas, other regions represent business zones.

Census and mobile phone information must be provided with respect to a common geographic grid to facilitate their use in emergency flight planning. This work uses the mobile phone grid as the common geographic grid. Therefore, census data are converted to the mobile phone grid.[§§] For this transformation, each mobile phone grid cell is assigned the fraction of the population given by the census based on the fraction of the overlapping area of both grid cells. This is the same approach used in [19]. The result is a census-based population value for each cell, $\lambda_{\text{census}}(g)$. The final preprocessing step is to normalize the population or occupancy values:

$$\hat{\lambda}_{\text{census}}(g) = \frac{\lambda_{\text{census}}(g)}{\sum_{g=1}^{n_g} \lambda_{\text{census}}(g)} \tag{12}$$

The normalized grid occupancy values then sum to 1.

### B.   OpenStreetMap

OpenStreetMap is an openly available database that contains information such as roads, buildings, and businesses. Although substantial data from government sources are currently available [1], the open database information surpasses "official" sources in several ways. For example, this paper evaluates the city of Milan, characterized in official open data[¶] and the Copernicus Earth Observation data.[***] The first official source

---

[††]"Linked Open Data," Instituto Nazionale di Statistica, Italy, http://datiopen.istat.it/dat/dataset/partizioniterritorialisubcomunali/Lombardia_Sezioni Censimento.zip [retrieved 20 June 2016].
[‡‡]Data available online at http://www.istat.it/storage/cartografia/basi_territoriali/WGS_84_UTM/2011/R03_11_WGS84.zip [retrieved 20 June 2016].
[§§]All GIS processing used in this work was performed with QGIS and additional plug-ins. QGIS is a standard and open-source GIS software.
[¶]"Open Data del Comune Milano," Comune di Milano, Italy, http://dati.comune.milano.it [retrieved 20 June 2016].
[***]"Copernicus — The European Earth Observation Programme," Copernicus Program, http://land.copernicus.eu [retrieved 20 June 2016].

provides different city data sets including, for example, school locations. Such locations are represented as points, whereas OpenStreetMap provides an outline of each building that appears to compare well with sample satellite photos. Copernicus includes CORINE Land Cover 2012, but OpenStreetMap has considerably more detail. OpenStreetMap vector fields can receive more than 900 labels. One shape can receive more than one label, such as a park that also has a fence. Shapes can overlap each other (e.g., a tree, represented by a node, can be inside a polygon that defines a park). Information regarding one region can be downloaded and filtered for interesting labels. Therefore, it is possible for example to build a database of all soccer fields in Milan.

## V.   Data Integration into Emergency Landing Planning

### A.   Emergency Landing Site Identification

Researchers have studied emergency landing site selection in previous work. One approach first identifies edge-free regions in real-time image data as possible landing sites and matches the results with map data [2,10]. Candidate landing areas are then evaluated for risk, and the best one is selected. A similar approach using lidar point-cloud data was proposed to optimize UAS landing site during a tracking mission [16]. An alternate database-centric approach follows the architecture for GIS data use proposed by [1], using a matrix or cell grid with associated cost to represent possible landing areas. The selected landing area may be a single cell [3] or a group of adjacent cells selected using a kernel [11].

This paper proposes the identification of candidate landing sites using database information only. A database-only strategy is required when visible areas are unsafe or the aircraft/UAS does not carry sensors capable of characterizing local area safety. Instead of computing a risk associated with each element in a grid [1], we propose infusion of GIS data in vector form. A candidate landing site can then be identified in an initial search for safe touchdown area; then, the risk to people on the ground, property damage, and aircraft damage can be factored into decision-making later. This approach is similar to that used by researchers relying on real-time sensor data. Selecting a landing site from vector data maintains all available information that can be used for subsequent risk computation. It also allows incorporation of real-time data from mobile phones if a data link is available.

Figure 9 presents the proposed data aggregation strategy separated into offline and online parts. The first offline preprocessing step is to identify data attributes indicating an area suitable for landing versus hazards to avoid. For example, land covered by crops is generally considered suitable for emergency landing, whereas electricity power lines are hazardous. Both information sets are extracted from the database. Hazardous elements are then grown in area to provide a safety margin. This work uses a fixed value for the applied area buffer, but this value could depend on type of hazard and height of buildings for example. The next step is to compare suitable and hazardous sites to eliminate any overlap between both. That can result in new subelements, for example, if a large crop area is traversed by a power line.

A large number of candidate landing sites can result from this process, only some of which are actually suitable for landing. Unsuitable examples include areas that are excessively small or narrow or that contain significant "gaps" that can represent obstacles. An automatic process to reliably eliminate such unsuitable areas maps is to restrict suitable areas to those that safely contain a circle or a rectangle of a specified minimum size. The associated mathematics is still an active research area, especially for nonconvex polygons with holes (e.g., Molano et al. [25]). Our work applies a simplified approach inspired by [10] and illustrated in Fig. 10.

The procedure starts from candidate landing site areas represented in Fig. 10a by the regions inside the two boundary lines. First, candidate landing areas are reduced using a buffer of negative distance. This distance corresponds to the radius of the minimum circular area acceptable as a landing site. This value is related to the size and landing dynamics of the particular aircraft. Fixed-wing aircraft, for example, would require a larger margin than a vertical-takeoff-and-landing aircraft. Particular failure cases might call for increased landing zone areas. This procedure
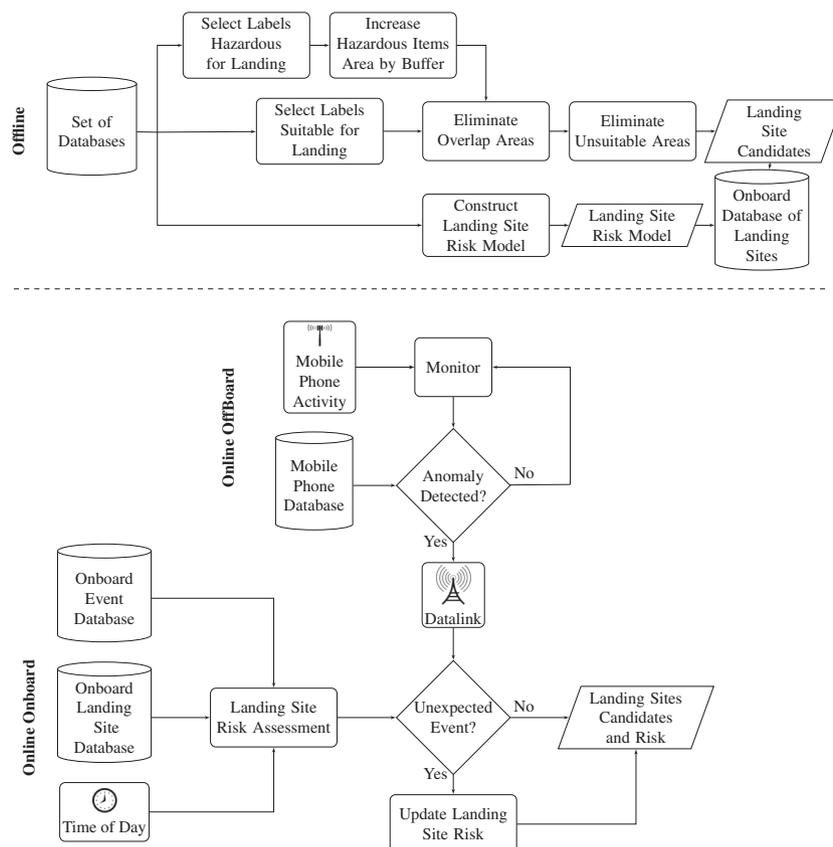


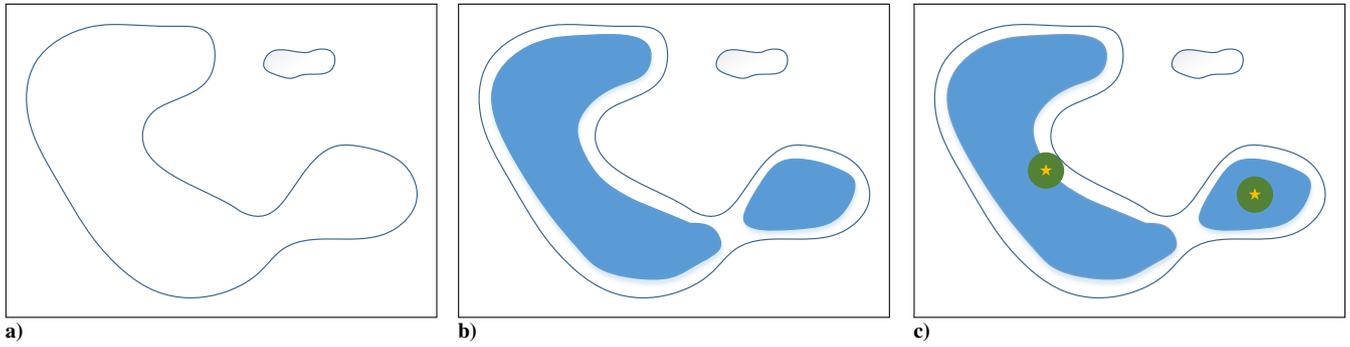**Fig. 9   Landing site selection from database fusion.**

**Fig. 10 Process of eliminating unsuitable landing sites.**

eliminates small landing areas and may divide one landing area into more than one region as illustrated in Fig. 10b. A "pseudocentroid" of each safe landing area is then computed restricting it to be within the contracted area boundary as required to ensure that it offers a sufficient, safe area. Figure 10c marks these pseudocentroids with stars. By construction, this point represents the center of a geographically suitable landing site with minimum size of a circle of radius equal to the applied buffer (depicted by a circle in Fig. 10c). Note that this process is conservative (i.e., the returned areas guarantee that minimum-area constraints are satisfied). However, it is not "optimal". For example, one region might support several different circular landing areas, but only one will be returned by the algorithm. Moreover, candidate areas may have a larger simply connected region in a location different from the one returned.

Although this work considered only automatic selection and classification of landing sites, an offline landing site identification process could benefit from human supervision and information fusion with other data sources. Although the number of possible landing sites can be large, human supervision may be necessary to build the required confidence in the data source and algorithm to accept automatic classification. Note that suitable landing sites do not need to be determined for each flight and UAS. A candidate landing site can be defined for a maximum UAS size or type for example. Regardless of how suitable sites are identified, the resulting areas can be considered candidate emergency landing sites.

For this work, all information contained in the GIS database for the landing area remains available even after landing site identification. For each suitable label/area, an emergency landing risk model is then constructed with three types of time-varying risk: risk to people (humans) on the ground ($R_h$), risk to property on the ground ($R_p$), and risk to the vehicle ($R_v$).[†††] The landing site risk model includes additionally a parameter describing landing area risk $R_a$. The total risk $R_{ls}$ associated with each candidate landing site $ls$ is given by a weighted sum:

$$R_{ls}(g, t) = \begin{bmatrix} W_h & W_p & W_v & W_a \end{bmatrix} \begin{bmatrix} R_h(g, t) \\ R_p(g, t) \\ R_v(g, t) \\ R_a(g) \end{bmatrix} \tag{13}$$

The final step of landing site database preprocessing is to connect the candidate landing site positions to the grid used for flight-path planning. This is done by associating each candidate landing area center to a cell of the path planner grid. Note that, because of the varying resolution of different datasets, one grid cell may have more than one candidate landing site, and so this study uses the simple approach of selecting the site in each grid with lowest total risk $R_{ls}$. The final output of landing site preprocessing is a set of candidate landing sites and associated risks:

$$G_{ls}(t) = \{(g_j, R_j(g_j, t)) | j = \{1, 2, \ldots, n_{ls}\}\} \tag{14}$$

where $n_{ls}$ is the total number of candidate landing sites. Before a flight, the UAS will be loaded with the candidate landing sites around its planned flight route. In the case of an in-flight emergency, the flight planner will use set $G_{ls}$ and the time $t$ to estimate the risk associated with each landing site classified as suitable based on GIS information. Planned events that are expected to attract an unusual occupancy concentration (e.g., a soccer match) can be loaded into the vehicle and used to increase the level of risk to people on the ground, although no specific event data interface is examined in this work.

Mobile phone activity discussed earlier can provide a means of estimating occupancy to supplement expected trends and scheduled event data. A cloud-based real-time system can monitor mobile phone activity and compare incoming data with historical trends. If an anomaly is observed, specific information can be sent to the aircraft via data link,[‡‡‡] and landing site risk can be reevaluated. Occupation anomalies (updates) can be found from a median model [9] for each landing site in cell $g$ given the current $(D, h_{10\,m})$:

$$\phi(g, t_{10\,m}) > (1 + k)\Phi(g, D, h_{10\,m}), t_{10\,m} \in D, t_{10\,m} \bmod 1440 = h_{10\,m} \Rightarrow \text{Anomaly} = \text{true} \tag{15}$$

An anomaly is detected every time the activity in one cell is greater than a factor $(1 + k)$ from the median. Higher values of $k$ require a greater amount of cell phone activity to be classified an anomaly, which increases the recall but may decrease precision [26]. Note that, because the mobile phone database offers data at 10 min intervals, the risk at each landing site is also updated at 10 min intervals. In addition, note that, distinct from [9], our work does not apply a separate threshold to select large events only. This work is instead interested in events that affect prior risk values due to people on the ground. Even if such events are small in terms of mobile phone activity, they can change the risk of a landing area. Moreover, in the case of planned events, an anomaly would only be only reported if the mobile phone activity indicates more than the expected number of people.

### B. Real-Time Occupancy Estimate

Census data provide information about where people live (i.e., where they usually sleep) but not where people travel during the day. Mobile phone activity can indicate where people are, but the level of activity varies significantly depending on the time of day. The ideal scenario for

---

[†††]This work assumes that an aircraft has declared an emergency and is in contact with air traffic control, and so no collision risk is modeled, given that the distressed aircraft has priority handling.

[‡‡‡]Such data link might also be made available through a mobile phone network.

occupancy estimation would be for every person to conduct exactly one mobile phone transaction during each time interval. In this case, occupancy would be directly proportional to mobile phone activity. In reality, not every person is using the phone; some are conducting many transactions per time interval, and data may not be available from all cellular service providers. Note that a presence data feature can provide a reasonable occupancy estimate so long as presence data are available across all carriers.

Independent of these challenges, it is reasonable to assume the accuracy of an occupancy measure from mobile phone activity increases with increased phone activity. This paper therefore proposes to combine normalized mobile phone aggregated SMS and call activity $\hat{\phi}$ and census occupancy $\hat{\lambda}_{\text{census}}$ to compute an occupancy estimate $\lambda_{rt}(g, t_{10\,\text{m}})$ for each grid cell $g$ and time $t$:

$$r_\lambda = \min\left[\frac{\phi_{\text{grid}}(t_{10\,\text{m}})}{\Theta(D)}, 1\right], \qquad t_{10\,\text{m}} \in D \tag{16}$$

$$\lambda_{rt}(g, t_{10\,\text{m}}) = (1 - r_\lambda)\hat{\lambda}_{\text{census}}(g) + r_\lambda\hat{\phi}(g, t_{10\,\text{m}}) \tag{17}$$

Here, $r_\lambda$ represents how the total activity on the grid compares to the usual maximum for that day of the week. Note that Eq. (16) simply distributes the weight given to the census and mobile phone estimates. Moreover, total grid occupancy is always normalized to 1, consistent with an assumption that the population inside the overall grid is constant.

Although Eq. (16) uses real-time mobile phone data, it can be approximated using the historical median if $\phi_{\text{grid}}(t_{10\,\text{m}})$ is substituted by $\Phi_{\text{grid}}(D, h_{10\,\text{m}})$ and $\hat{\phi}(g, t_{10\,\text{m}})$ is substituted by $\hat{\Phi}(D, h_{10\,\text{m}})$ with $t_{10\,\text{m}} \in D$ and $h_{10\,\text{m}} = t_{10\,\text{m}}(\text{mod}\,1440)$. With this procedure, aircraft without data link are able to use information stored before a flight in onboard databases.

## C. Data Use in Emergency Path Planning

The candidate landing site, its risks, and occupancy data can be used by a flight planner to compute an emergency landing path or trajectory. A variety of motion planning algorithms has been proposed for UAS. A 2009 survey of motion planners applicable to UAS is presented in [27]. Examples of more-recent motion planners include RRT-AR* [28], fast-marching squared [29], semi-Lagrangian schemes [30], and fast-marching trees [31]. Although these newer methods may be considered more suitable for continuous configuration spaces and may be able to handle differential constraints, this paper makes use of conventional A* algorithm to illustrate the use of the candidate landing site risk assessment and occupancy estimates proposed in this article. This information can be similarly employed for different motion planners by considering occupancy estimates and landing site risks in motion planner cost functions.

Consider an application of occupancy data to a two-dimensional path planning algorithm applicable to a small UAS transiting over a city. A conventional A* algorithm is applied, and the search space corresponds to the mobile phone grid. This search space is increased with additional states associated with each landing site (i.e., there are two possible states associated with each grid cell $g$ containing a candidate landing site: flying or landed). Goal nodes correspond to landing states of each grid position with a candidate landing site. This work assumes that the small UAS is flying at a constant altitude above buildings, and thus no obstacles are considered.

The A* search graph assumes that the aircraft (small UAS) can only move to eight adjacent cells (north, northeast, east, southeast, south, southwest, west, or northwest), or the aircraft can land if it is over a candidate landing site grid. Actual transition cost from cell $g$ to cell $g'$ when selecting action $a$ is given by

$$C(g, a, g') = \begin{cases} w\left[\frac{1}{2}\lambda_{rt}(g, t_{10\,\text{m}}) + \frac{1}{2}\lambda_{rt}(g', t_{10\,\text{m}}) + \lambda_0\right] & a \in \{N, E, S, W\} \\ \sqrt{2}w\left[\frac{1}{2}\lambda_{rt}(g, t_{10\,\text{m}}) + \frac{1}{2}\lambda_{rt}(g', t_{10\,\text{m}}) + \lambda_0\right] & a \in \{NE, SE, SW, NW\} \\ R_{ls}(g, t) & a \in \{\text{Land}\}, g \in G_{ls} \end{cases} \tag{18}$$

In the case of routine transit to an adjacent flight grid, the cost of action $a$ is equal to the product of a weight $w$ and the sum of real-time occupancy of both traversed cells $g$ and $g'$ and a transition cost $\lambda_0$. Note that, in the case of diagonal movement, the cost is scaled proportionally to the increase in flight distance relative to a nondiagonal transit. Weight $w$ can be set to adjust prioritization of overflight versus landing risk. For transitions to a landing state, the cost is equal to the landing site risk $R_{ls}(g, t)$.

The A* search node heuristic value is computed assuming that real-time occupancy is zero for all cells, then using the minimum-cost value computed over all candidate landing sites. The heuristic considers a feasible flight path given the search grid structure and available actions. Specifically, it assumes that the aircraft moves in a diagonal line until reaching the same row or column of the landing site and then proceeds in a nondiagonal path to the landing site:

$$d_j(g, g_j) = \left|d_{\text{rows}}(g, g_j) - d_{\text{cols}}(g, g_j)\right| \tag{19}$$

$$h(g) = \min_{g_j \in G_{ls}} \left\{w\lambda_0\left[d_j + \sqrt{2}\left|d_{\text{rows}}(g, g_j) - d_j(g, g_j)\right|\right] + R_{ls}(g_j, t)\right\} \tag{20}$$

In the preceding expression, $d_{\text{rows}}(g, g_j)$ and $d_{\text{cols}}(g, g_j)$ are the number of cells in the column and row directions between the current cell and a candidate landing site cell, respectively. This heuristic is consistent and thus assures optimality of each A* graph search.

## VI.  Case Study

### A.  Emergency Landing Site Identification

OpenStreetMap labels presented in the Milan data set were investigated for landing site suitability. Following basic pilot knowledge, the suitable landing site search goal is to identify open spaces with few obstacles. The resulting labels are collected into groups that pose similar risks during an emergency landing. Results of this investigation are presented in Table 3. This table illustrates the information currently available from OpenStreetMap.

Map labels selected in Table 3 do not include all that could be suitable for an emergency landing. Two main simplifications were used. First, only "polygons" or "multipolygons" were used in this work as suitable landing areas. Smaller rivers, for example, are mapped as lines with

**Table 3  Selected and grouped OpenStreetMap labels classified as suitable for emergency landing**

| Key | Value | Description |
|---|---|---|
| | | *Group 1: Parking lot* |
| Amenity | Parking | Car parking lot |
| | | *Group 2: Grass* |
| Land use | Grass | Areas covered with grass |
| Land use | Village green | An open green space for general recreation; may include pitch, nets, and so on; usually municipal but possibly also private colleges or companies |
| Land use | Recreation ground | An area of common land, usually grass, in the center of a village |
| | | *Group 3: Greenfield* |
| Land use | Greenfield | Land scheduled for new development where were previously no buildings; represents a greenfield scheduled to become a construction site |
| | | *Group 4: Brownfield* |
| Land use | Brownfield | Land scheduled for new development where old buildings have been demolished and cleared; a brownfield that is scheduled to turn into a construction site |
| | | *Group 5: Meadow* |
| Natural | Meadow | An area of land primarily vegetated by grass and other nonwoody plants, usually mowed for hay |
| Natural | Grassland | Land where vegetation is dominated by grasses and other herbaceous (nonwoody) plants, except for ornamental grass, mowing for hay, etc. |
| Natural | Heath | Bare lower-lying uncultivated land with bushes but little or no tree cover |
| Leisure | Garden | Location where flowers and other plants are grown in a decorative and structured manner or for scientific purposes |
| | | *Group 6: Sports ground* |
| Leisure | Pitch | A field for playing football/soccer, cricket, baseball-style sports, and skate parks |
| Leisure | Golf course | Land hosting a golf course |
| Leisure | Stadium | A major sports arena with substantial tiered seating |
| Leisure | Track | Running, cycle-racing, greyhound, horses, etc. |
| | | *Group 7: Water* |
| Natural | Water | Lakes, etc. |
| Waterway | Riverbank | Designation of larger rivers and waterways; can define an area between opposite riverbanks |
| | | *Group 8: Farmland* |
| Land use | Farmland | An area of farmland used for tillage and pasture (animals, crops, vegetables, flowers, fruit growing) |

**Table 4  Selected and grouped OpenStreetMap labels classified as hazardous for emergency landing**

| Key | Value | Description |
|---|---|---|
| Building | All | Cells containing individual buildings or groups of connected buildings |
| Power | All but cable | Electrical power generation and distributions systems; value "cable" is used in the database to describe underground or underwater cables; overground lines are represented by the value "lines" |
| Man-made | All | Man-made (artificial) structures that are added to the landscape |
| Highway | All but tunnel | Road, street, or path |
| Railway | All but tunnel | Railways, including mainline services, subways, heritage lines, and trams |
| Natural | Tree | Lone or clusters of trees |
| | Tree row | Line of trees |
| | Rock | A notable rock or group of rocks |
| | Stone | Freestanding stone (e.g., glacial erratic) |

information about cell width. Use of these regions would require additional preprocessing to compute the landing area and thus is left for future work. The same applies to any kind of traffic element such as roads or rails, which were considered as obstacles and not potential landing sites. Second, any airfield or helipad area is not considered in this work as a landing site candidate, although grass areas close to runways were automatically included by the proposed method. Note that some areas may have two labels overlapping; in this case one unique label for each area is selected.

Even though the labels in Table 3 represent areas generally free of obstacles, certain obstructions may still be present. Power lines, for example, can traverse open fields, and a sports ground can be covered with a partial or full roof. Table 4 presents map labels used to identify things that could be hazardous for a landing. In this case, any map element type is used, including points and lines. Highways and railways were included because they can overlay farm, parks, and sports ground. Although not exactly an obstacle, the presence of a path or a track (both mapped under the highway tag) may increase the chance of nonmapped small obstacles such as light poles and fences. Using highways as emergency landing sites was considered in different work [5].

Figure 11 presents the hazard layer considered at a small part of the interest region west of downtown Milan. A 10 m buffer is applied to all hazards. The identified landing sites in the same region are presented in Fig. 12. Different map labels are presented in different colors, and the result of a 30 m negative buffer is visible by the darker areas. Each computed centroid is marked with a star. Figure 12 also shows the mobile phone database grid that is used for path planning. Note that one grid cell may offer more than one landing site due to granularity differences between the different databases.

An example risk model based on landing site groups and time of day is presented in Table 5. Characteristics of each landing site group are combined with UAS properties and landing requirements to classify UAS landing risk at different times of the day. For example "sports grounds"
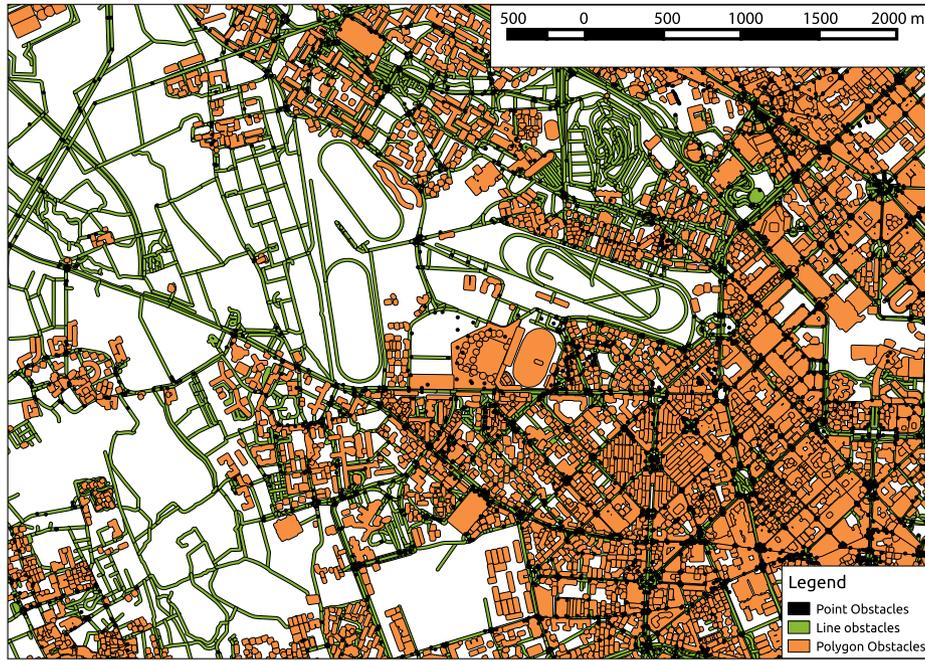
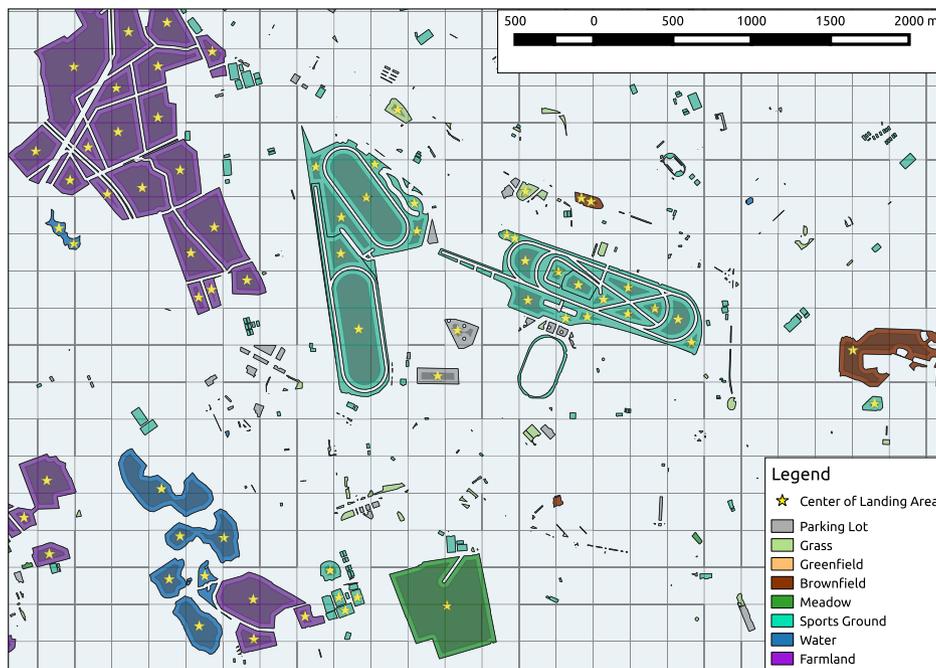**Fig. 11    Hazard map features with a buffer of 10 m.**



**Fig. 12    Candidate landing site areas. Colors indicate groups from Table 3.**

are more often occupied by people at night but do not often contain valuable property and do not pose a significant risk to the vehicle. Landing on "water" reduces the chances of harm to people and property on the ground, but a water landing is likely to damage most contemporary UAS. Risk values are abstracted to scales of low, medium, and high for illustration. A careful discussion with stakeholders would be required to establish acceptable risk levels. When a significant event increasing occupancy is detected, the risk to a human on the ground is elevated one level. For this case study, the following risk levels were defined:

$$\text{Risk} = \{\text{Low, Medium, High}\} = \{0, 0.5, 1\} \tag{21}$$

The reduced area after the application of the negative buffer illustrated in Fig. 9 is used as a measurement of the risk associated with the landing area $R_a$. These reduced areas were computed for all identified landing sites, and 20, 40, 60, and 80% quantiles were determined. For each candidate landing site then, $R_a \in \{0, 0.25, 0.5, 0.75, 1.0\}$ in an inverse relation to the quantile of its reduced area (i.e., large landing areas corresponds to lower $R_a$ values).

In this case study, the landing site final risk is computed using a uniform distribution of weights:

$$W_p = W_h = W_v = W_a = 0.25 \tag{22}$$

**Table 5    Landing risk values for landing site groups**

| Group | Description | Time of day | Risk to Humans $R_h$ | Property $R_p$ | Vehicle $R_v$ |
|---|---|---|---|---|---|
| 1 | Parking lot | Off-hours | Medium | Medium | Medium |
| | | Business hours | Medium | High | High |
| 2 | Grass | Night (any) | Low | Low | Low |
| | | Off-hours | High | Low | Low |
| | | Business hours | Medium | Low | Low |
| 3 | Greenfield | Any | Low | Low | Low |
| 4 | Brownfield | Any | Low | Low | Low |
| 5 | Meadow | Night | Low | Low | Medium |
| | | Off-business | Medium | Low | Medium |
| | | Business | Medium | Low | Medium |
| 6 | Sports ground | Night | Medium | Low | Low |
| | | Off-business | High | Low | Low |
| | | Business | High | Low | Low |
| 7 | Water | Any | Low | Low | High |
| 8 | Farm | Any | Low | Low | Medium |

For event detection, this case study uses $k = 1$ (i.e., an event is detected if the mobile activity in a grid cell is greater than twice the median for that cell and time of day). Figure 13 presents the histogram of computed risk of candidate landing sites for 4th December 2013 at 08:30 p.m. Because of risk model characteristics, risk values are distributed in a discrete set (i.e., all landing sites represented by each bar in Fig. 13 have the exact same risk). Note the significant concentration of landing sites with intermediate risk values.

Figure 14 presents the cells that contain candidate landing sites overlaying the Milan map. There is a clear contrast between downtown and city outskirts in the number of identified landing sites. The black rectangle represents approximately the area presented in Figs. 11 and 12. The connection between candidate landing site position and their assigned grid cell can be observed when comparing the star distribution and grid in Fig. 12 to highlighted cells in Fig. 14.

Cells in Fig. 14 are colored based on computed landing risk for the same day and hour of Fig. 13. Higher risk cells are colored dark red. Candidate landing sites are divided between selected colors based on its quantile. Candidate landing site whose computed risk was affected by an event detected by mobile phone activity are presented in Fig. 14 by the lack of a black border around the cell. The diamond in the same figure mark the position of San Siro Stadium and the selected date and time corresponds to a soccer game. Note that an event was detected in the cells close to the stadium as expected.

### B. Flight Planning

This case study uses a path cost $\lambda_0 = 5.10^{-5}$ (i.e., half of the occupancy value if it was uniformly distributed over the grid). Weight $w$ is used to balance the importance given to landing site risk and path risk. Moreover, it allows for different scales to be used by both risk measurements. Landing site risk goes from 0 to 1, whereas the transition risk based on occupancy has an order of magnitude $10^{-3}$. A low value of $w$ results in more importance given to the landing site risk. Resulting paths would always go to the closest candidate landing site from the group with the minimum value of landing risk (see Fig. 13). A high value of $w$ results in paths that mostly go to the closest candidate landing site independent of its landing risk. This case study uses a value of $w = 50$, resulting in balance and different results during the course of a weekday. Different values of $w$ would
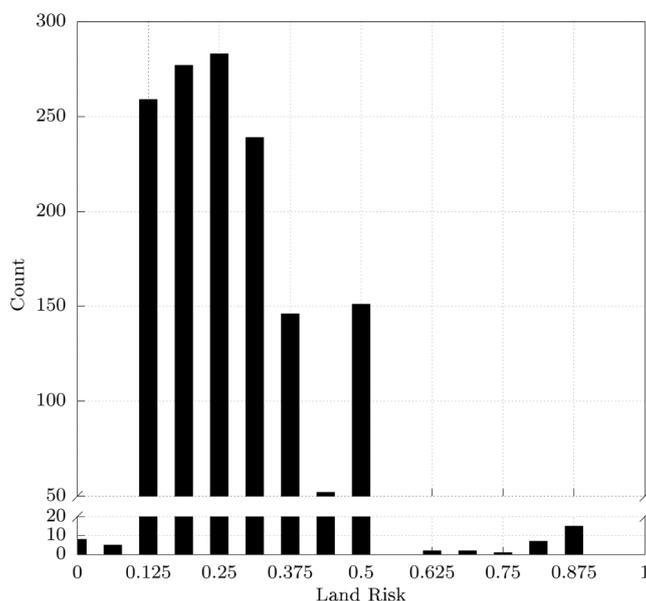


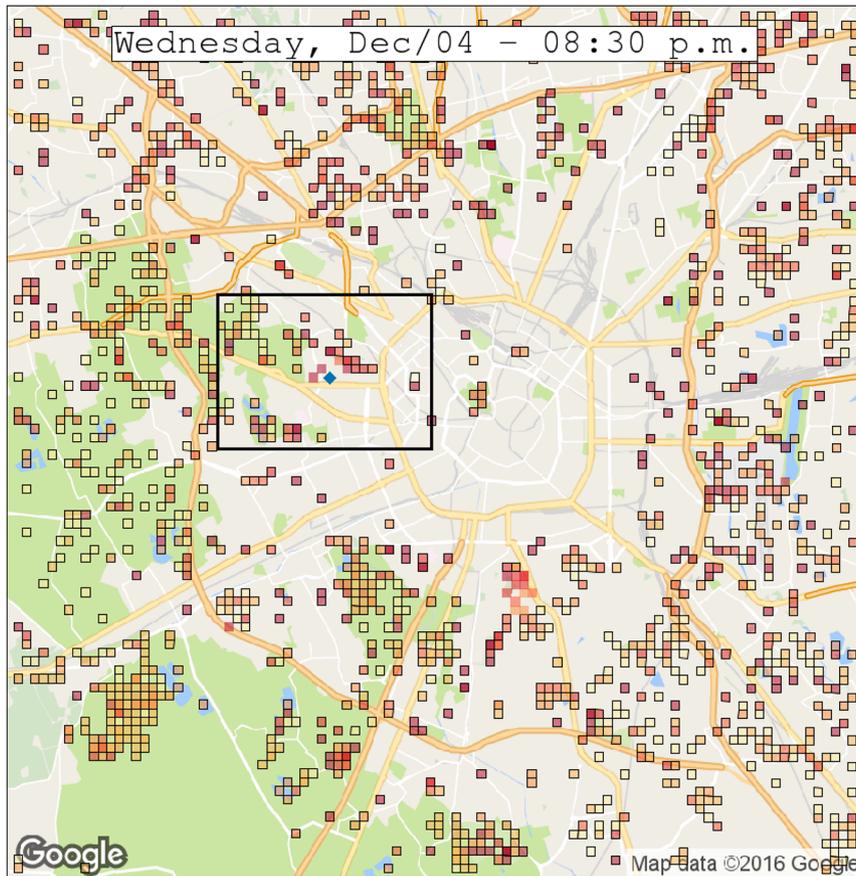**Fig. 13    Histogram of computed landing site risk on 4th December 2013, 08:30 p.m.**

**Fig. 14  Cells with candidate landing site areas colored according to risk.**

be suitable for different failure modes. For example, a low-energy scenario requires a near-term landing, and so $w$ should be high, whereas a wing-flap failure for a fixed-wing UAS should adopt a small value of $w$ because flaps are most critical for landing.

Figures 15 and 16 illustrate path planning solutions for different hours of a Friday morning. Note that these figures only present a part of the downtown area included by Fig. 14. Paths are presented over a grid colored in accordance with estimated real-time occupancy. Colors are divided based in quantiles, and so the thresholds of each figure can vary. Nevertheless, the two figures show that the occupancy estimation method produces logical results. People are expected to be at home early morning, and occupancy follows census data. During business hours, there is increased occupancy in downtown, estimated primarily from mobile phone activity.

The initial state of the small UAS is indicated by a triangle. It is selected to be in a downtown region far from any candidate landing site. The chosen landing site in each solution is indicated by a circle in each figure, whereas the selected emergency flight path is represented by a solid line. For both cases, a "direct" path from the initial state to landing site is included for comparison. The direct path respects the search grid pattern as given in the Eq. (20) heuristic.

During the early morning (Fig. 15), there are few people in downtown, and so the risk-optimal path is to transit over the downtown region and land on the large brownfield that appears in Fig. 12. Note that landing site cost indicated in the figure legend is zero due to the low risk of the chosen site and our choice of risk parameters in Sec. VI.A. The computed best path avoids overflying the areas with higher risk as expected. Figure 17 shows the same paths as Fig. 15, but instead of estimated occupancy, the underlined cells represent candidate landing sites as per Fig. 14. Note that the small UAS flies close to multiple candidate landing sites but selects the one with a lower risk.

After the business hours begin, downtown has high occupancy, and so the planner returns a path that exits downtown to the south and lands in a farm field (Fig. 16), which represents a higher landing risk than a brownfield but a lower overall path cost. These results illustrate the tradeoff performed by the flight planner in overflight versus landing risk as well as the solution differences that might occur over a normal weekday.

## VII.  Discussion

This work builds a landing flight plan based on occupancy observed to vary over time of day. Several improvements are possible in future work. First, other data such as terrain slope and building heights should be integrated for landing site selection and flight planning, respectively. Obstacles or no-fly areas must also be taken into account by the flight planner. Second, sensor–database fusion is critical for the final landing approach per the model presented in Fig. 1. Third, for a real application, the different risk assessments made throughout the paper must be tailored for a particular UAS and its failure modes.

Additional research is required to verify and improve the proposed occupancy model and associated cost metrics proposed in this paper. As discussed in Sec. II.B, a variety of sensors and data processing methods could be used to further improve occupancy information from mobile phone reports. These methods provide different space and time resolution as well as accuracy. This paper uses mobile-phone activity in a manner that can characterize higher-than-usual occupancy without precisely defining the number of people or their location within a grid. Such data would, therefore, need to be supplemented with onboard sensor data on approach to landing. On the other hand, occupancy grids used for flight planning do not require high precision, and so mobile phone activity combined with census estimates will likely be adequate for use in autonomous flight planning applications. Note that, although cell phone data can be sensitive and proprietary, there is at least one company in the United States whose main business is to provide occupancy estimates based on mobile phone data provided by different carriers [32].
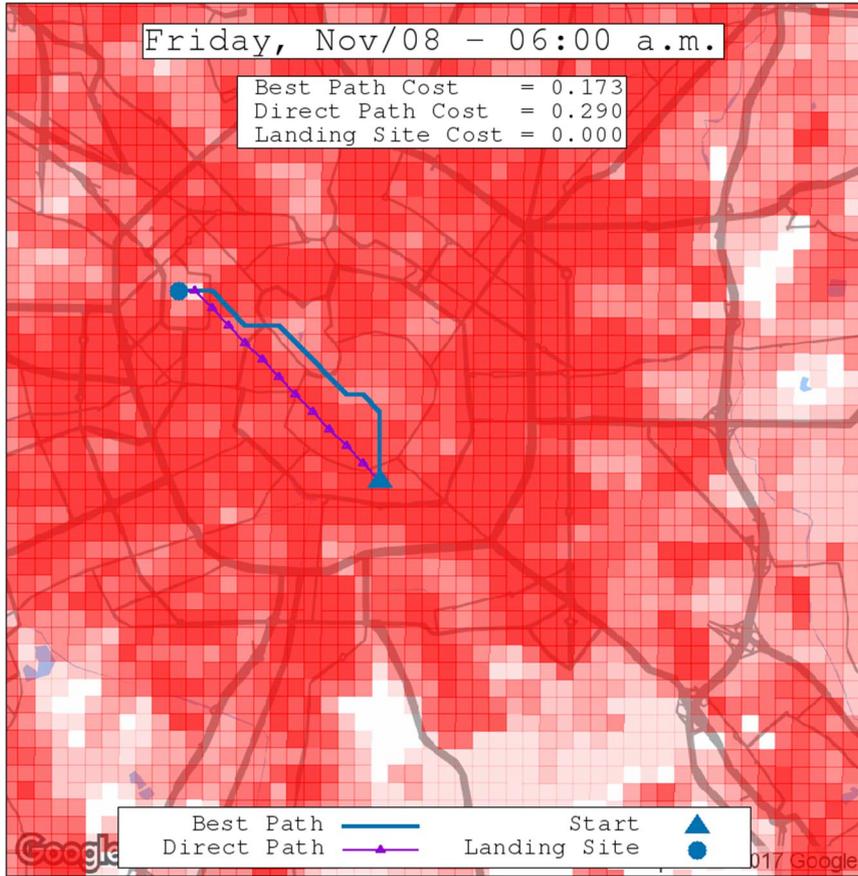
**Fig. 15    Occupancy map and planned emergency landing path: 8th November 2013, 6:00 a.m.**
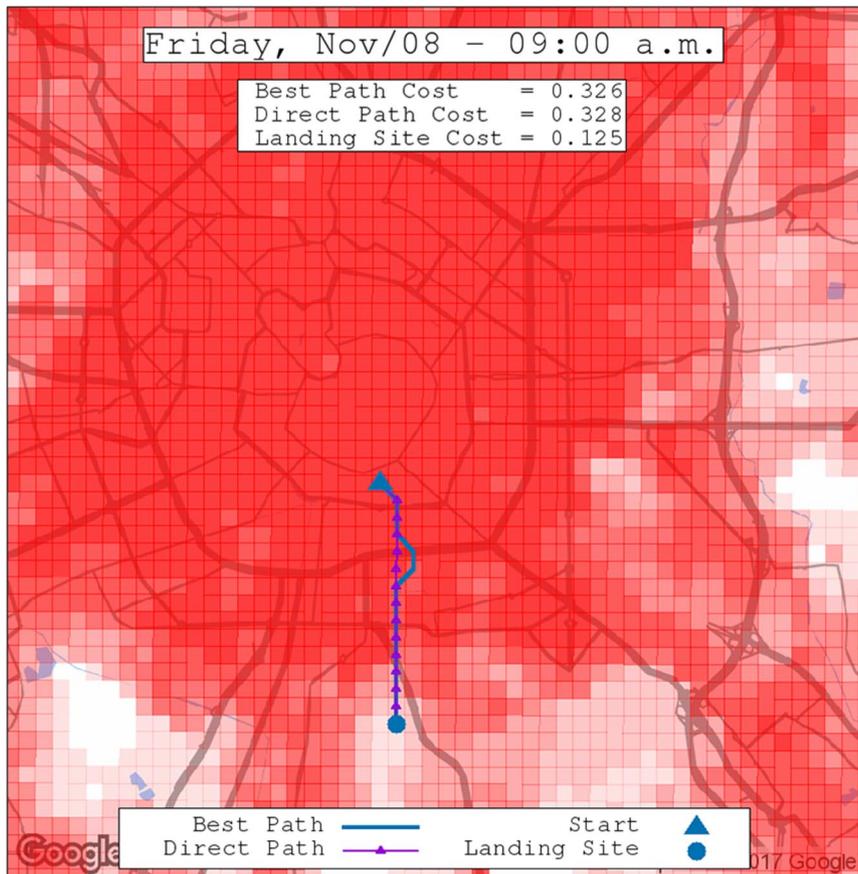


**Fig. 16    Occupancy map and planned emergency landing path: 8th November 2013, 9:00 a.m.**
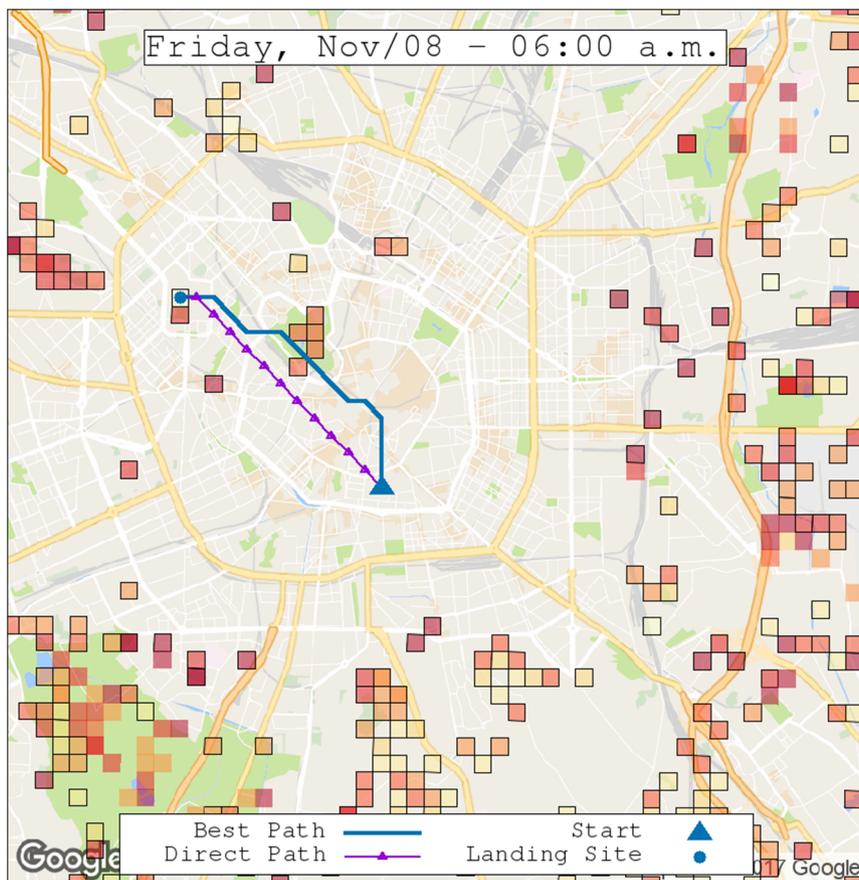
**Fig. 17    Candidate landing sites and planned emergency landing path: 8th November 2013, 6:00 a.m.**

A two-dimensional flight planner was considered in this paper because the focus was the inclusion of risk to people on the ground as an small UAS overflies a region at constant altitude. With the extension to three dimensions, height above ground might also scale risk to people on ground. As discussed in Sec. V.C, more-advanced motions planners can be employed to provide lower-cost paths over a continuous state space that considers differential constraints. Furthermore, the flight planner proposed in this paper is static (i.e., a flight plan is determined based on an initial occupancy estimate and then executed to completion). A small UAS with a cruise speed of 50 mph can cross one grid cell in approximately 10.5 s, whereas mobile data used in this work are updated in 10 min intervals. This would allow the small UAS to cross 57 cells before new data are available, supporting the single-plan generation strategy in this work. However, path planning could also use historical or recent dynamics in mobile phone data to improve future occupancy predictions. Updates to flight plans based on unexpected occupancy changes can also be made.

Finally, although this paper is focused on emergency landing, detect-and-avoid is also a critical challenge for large-scale UAS deployment in urban areas. A data link can significantly simplify this challenge through direct high-speed communication and coordination across all traffic. Data link also offers emergency flight planners the ability to take into account other traffic to minimize the impact of the emergency path on other UAS along with rapidly alerting other UAS of the emergency and associated landing plan.

## VIII.    Conclusions

This paper proposes the infusion of new data sources to guide an emergency landing planner to overfly regions and land at sites with low risk to property, people on the ground, and the aircraft itself. The proposed method extracts candidate landing sites from a geographic information system database, taking advantage of ground usage information to estimate landing risk, dependent on the time of the day. Publicly available mobile phone data are combined with census data to provide real-time occupancy estimates. Landing site risk based on land use is combined with mobile phone data to assess the overall risk posed by the aircraft to people on the ground. A flight planning case study in the Milan, Italy, region is presented that considers both landing site risk and overflight risk computed from occupancy estimates.

## Acknowledgments

## References

[1]  Rackliffe, N., Yanco, H. A., and Casper, J., "Using Geographic Information Systems (GIS) for UAV Landings and UGV Navigation," *Proceedings of the 2011 IEEE Conference on Technologies for Practical Robot Applications (TePRA)*, IEEE Publ., Piscataway, NJ, 2011, pp. 145–150.
    doi:10.1109/TEPRA.2011.5753497
[2]  Mejias, L., Fitzgerald, D. L., Eng, P. C., and Xi, L., "Forced Landing Technologies for Unmanned Aerial Vehicles: Towards Safer Operations," *Aerial Vehicles*, edited by Lam, T. M., In-Tech, Rijeka, Croatia, 2009, Chap. 21.
    doi:10.5772/6481

[3] Ten Harmsel, A. J., Olson, I. J., and Atkins, E. M., "Emergency Flight Planning for an Energy-Constrained Multicopter," *Journal of Intelligent and Robotic Systems*, Vol. 85, No. 1, Jan. 2017, pp. 1–21.
doi:10.1007/s10846-016-0370-z

[4] Chen, M., Hu, Q., Fisac, J., Akametalu, K., Mackin, C., and Tomlin, C., "Reachability-Based Safety and Goal Satisfaction of Unmanned Aerial Platoons on Air Highways," *Journal of Guidance, Control, and Dynamics*, 2016.
doi:10.2514/1.G000774

[5] Di Donato, P. F., and Atkins, E. M., "An Off-Runway Emergency Landing Aid for a Small Aircraft Experiencing Loss of Thrust," *AIAA Infotech @ Aerospace*, AIAA Paper 2015-1798, 2015.
doi:10.2514/6.2015-1798

[6] Naboulsi, D., Fiore, M., Ribot, S., and Stanica, R., "Large-Scale Mobile Traffic Analysis: A Survey," *IEEE Communications Surveys & Tutorials*, Vol. 18, No. 1, 2016, pp. 124–161.
doi:10.1109/COMST.2015.2491361

[7] Khodabandelou, G., Gauthier, V., El-Yacoubi, M., and Fiore, M., "Population Estimation from Mobile Network Traffic Metadata," *Proceedings of the 2016 IEEE 17th International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, IEEE Publ., Piscataway, NJ, 2016, pp. 1–9.
doi:10.1109/WoWMoM.2016.7523554

[8] Botta, F., Moat, H. S., and Preis, T., "Quantifying Crowd Size with Mobile Phone and Twitter Data," *Royal Society Open Science*, Vol. 2, No. 5, 2015, Paper 150162.
doi:10.1098/rsos.150162

[9] Cecaj, A., and Mamei, M., "Data Fusion for City Life Event Detection," *Journal of Ambient Intelligence and Humanized Computing*, Vol. 8, No. 1, 2017, pp. 117–131.
doi:10.1007/s12652-016-0354-7

[10] Patterson, T., McClean, S., Morrow, P., Parr, G., and Luo, C., "Timely Autonomous Identification of UAV Safe Landing Zones," *Image and Vision Computing*, Vol. 32, No. 9, 2014, pp. 568–578.
doi:10.1016/j.imavis.2014.06.006

[11] Bleier, M., Settele, F., Krauss, M., Knoll, A., and Schilling, K., "Risk Assessment of Flight Paths for Automatic Emergency Parachute Deployment in UAVs," *IFAC-PapersOnLine*, Vol. 48, No. 9, 2015, pp. 180–185.
doi:10.1016/j.ifacol.2015.08.080

[12] Atkins, E. M., Portillo, I. A., and Strube, M. J., "Emergency Flight Planning Applied to Total Loss of Thrust," *Journal of Aircraft*, Vol. 43, No. 4, 2006, pp. 1205–1216.
doi:10.2514/1.18816

[13] Meuleau, N., Plaunt, C., Smith, D. E., and Smith, T. B., "An Emergency Landing Planner for Damaged Aircraft," *Proceedings of the 21st Innovative Applications of Artificial Intelligence Conference*, edited by Haigh, K., and Rychtyckjy, N., AAAI Press, Menlo Park, CA, 2009, pp. 114–121.

[14] Cesetti, A., Frontoni, E., Mancini, A., Zingaretti, P., and Longhi, S., "A Vision-Based Guidance System for UAV Navigation and Safe Landing Using Natural Landmarks," *Journal of Intelligent and Robotic Systems*, Vol. 57, No. 1, Jan. 2010, pp. 233–257.
doi:10.1007/s10846-009-9373-3

[15] Cesetti, A., Frontoni, E., Mancini, A., and Zingaretti, P., "Autonomous Safe Landing of a Vision Guided Helicopter," *Proceedings of the 2010 IEEE/ASME International Conference on Mechatronics and Embedded Systems and Applications (MESA)*, IEEE Publ., Piscataway, NJ, July 2010, pp. 125–130.
doi:10.1109/MESA.2010.5552081

[16] Mackay, J., Ellingson, G., and McLain, T. W., "Landing Zone Determination for Autonomous Rotorcraft in Surveillance Applications," *AIAA Guidance, Navigation, and Control Conference*, AIAA Paper 2016-1137, 2016.
doi:10.2514/6.2016-1137

[17] Idicula, J., Akametalu, K., Chen, M., Tomlin, C., Ding, J., and Hook, L., "Where to Land: A Reachability Based Forced Landing Algorithm for Aircraft Engine Out Scenarios," NASA Document 20150023020, Nov. 2015, http://hdl.handle.net/2060/20150023020 [retrieved 30 Jan. 2017].

[18] Rufa, J. R., and Atkins, E. M., "Unmanned Aircraft System Navigation in the Urban Environment: A Systems Analysis," *Journal of Aerospace Information Systems*, Vol. 13, No. 4, April 2016, pp. 143–160.
doi:10.2514/1.I010280

[19] Douglass, R. W., Meyer, D. A., Ram, M., Rideout, D., and Song, D., "High Resolution Population Estimates from Telecommunications Data," *EPJ Data Science*, Vol. 4, No. 4, 2015, pp. 1–13.
doi:10.1140/epjds/s13688-015-0040-6

[20] Xavier, F. H. Z., Silveira, L. M., Almeida, J. M. d., Ziviani, A., Malab, C. H. S., and Marques-Neto, H. T., "Analyzing the Workload Dynamics of a Mobile Phone Network in Large Scale Events," *Proceedings of the 1st Workshop on Urban Networking, UrbaNe '12*, Association for Computing Machinery, New York, 2012, pp. 37–42.
doi:10.1145/2413236.2413245

[21] Bagrow, J. P., Wang, D., and Barabási, A.-L., "Collective Response of Human Populations to Large-Scale Emergencies," *PLOS ONE*, Vol. 6, No. 3, March 2011, Paper e17680. 10.1371/journal.pone.0017680

[22] Isaacman, S., Becker, R., Cáceres, R., Martonosi, M., Rowland, J., Varshavsky, A., and Willinger, W., "Human Mobility Modeling at Metropolitan Scales," *Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services, MobiSys '12*, Association for Computing Machinery, New York, 2012, pp. 239–252.
doi:10.1145/2307636.2307659

[23] Furno, A., Stanica, R., and Fiore, M., "A Comparative Evaluation of Urban Fabric Detection Techniques Based on Mobile Traffic Data," *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015, ASONAM '15*, Association for Computing Machinery, New York, 2015, pp. 689–696.
doi:10.1145/2808797.2810057

[24] Leys, C., Ley, C., Klein, O., Bernard, P., and Licata, L., "Detecting Outliers: Do Not Use Standard Deviation Around the Mean, Use Absolute Deviation Around the Median," *Journal of Experimental Social Psychology*, Vol. 49, No. 4, 2013, pp. 764–766.
doi:10.1016/j.jesp.2013.03.013

[25] Molano, R., Rodrguez, P. G., Caro, A., and Durán, M. L., "Finding the Largest Area Rectangle of Arbitrary Orientation in a Closed Contour," *Applied Mathematics and Computation*, Vol. 218, No. 19, 2012, pp. 9866–9874.
doi:10.1016/j.amc.2012.03.063

[26] Ting, K. M., "Precision and Recall," *Encyclopedia of Machine Learning*, edited by Sammut, C., and Webb, G. I., Springer, Boston, MA, 2010, pp. 781–781.
doi:10.1007/978-0-387-30164-8_652

[27] Goerzen, C., Kong, Z., and Mettler, B., "A Survey of Motion Planning Algorithms from the Perspective of Autonomous UAV Guidance," *Journal of Intelligent and Robotic Systems*, Vol. 57, No. 1, 2010, pp. 65–100.
doi:10.1007/s10846-009-9383-1

[28] Choudhury, S., Scherer, S., and Singh, S., "RRT*-AR: Sampling-Based Alternate Routes Planning with Applications to Autonomous Emergency Landing of a Helicopter," *Proceedings of the 2013 IEEE International Conference on Robotics and Automation*, IEEE Publ., Piscataway, NJ, May 2013, pp. 3947–3952.
doi:10.1109/ICRA.2013.6631133

[29] Gonzlez, V., Monje, C. A., Moreno, L., and Balaguer, C., "Fast Marching Square Method for {UAVs} Mission Planning with Consideration of Dubins Model Constraints," *IFAC-PapersOnLine*, Vol. 49, No. 17, 2016, pp. 164–169.
doi:10.1016/j.ifacol.2016.09.029

[30] Takei, R., and Tsai, R., "Optimal Trajectories of Curvature Constrained Motion in the Hamilton–Jacobi Formulation," *Journal of Scientific Computing*, Vol. 54, No. 2, 2013, pp. 622–644.
doi:10.1007/s10915-012-9671-y

[31] Schmerling, E., Janson, L., and Pavone, M., "Optimal Sampling-Based Motion Planning Under Differential Constraints: The Driftless Case," *Proceedings of the 2015 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE Publ., Piscataway, NJ, May 2015, pp. 2368–2375.
doi:10.1109/ICRA.2015.7139514

[32] Birdsall, M., "AirSage and ITE: Partnering for the Future of Transportation," *ITE Journal*, Vol. 83, No. 11, Nov. 2013, pp. 16–17.

M. J. Kochenderfer
*Associate Editor*