Statistical Characteristics of Convective Storms in Belgium Derived from Volumetric Weather Radar Observations

EDOUARD GOUDENHOOFDT AND LAURENT DELOBBE
Royal Meteorological Institute, Brussels, Belgium

(Manuscript received 21 March 2012, in final form 8 November 2012)

ABSTRACT
High-resolution volumetric reflectivity measurements from a C-band weather radar are used to study the characteristics of convective storms in Belgium. After clutter filtering, the data are processed by the storm-tracking system Thunderstorm Identification, Tracking, Analysis, and Nowcasting (TITAN) using a 40-dBZ reflectivity threshold. The 10-yr period of 5-min data includes more than 1 million identified storms, mostly organized in clusters. A storm is observed at a given point 6 h yr\(^{-1}\) on average. Regions of slightly higher probability are generally correlated with orographic variations. The probability of at least one storm in the study area is 15%, with a maximum of 35% for July and August. The number of storms, their coverage, and their water mass are limited most of the time. The probability to observe a high number of storms reaches a maximum in June and in the early afternoon in phase with solar heating. The probability of large storm coverage and large water mass is highest in July and in the late afternoon. Convective storms are mostly small and weak. Deeper ones are found mainly in the afternoon whereas bigger and more intense ones also appear in the evening. The occurrence of the most intense storms does not vary along the day. Simple tracks have a mean duration of 25 min. Complex tracks, involving splitting or merging, last 70 min on average. Most convective storms move in the northeast direction, with a median speed of 30 km h\(^{-1}\). Their motion is slower in summer and in the afternoon. Regions with slightly higher convective initiation are related to orography.

1. Introduction
The goal of this paper is to provide a comprehensive statistical analysis of convective storms in and around Belgium. The area of interest is characterized by a temperate climate, a nearby coastline, and intermediate level of orography (maximum 694 m). Convective storms are common meteorological phenomena that involve complex processes at different temporal and spatial scales. Severe events cause flash floods, strong winds, hail falls, and tornadoes, which can significantly affect human activities. Because of its complexity, the nowcasting (i.e., very short-range forecasting) of convective storms initiation and development remains a challenging problem. Despite major progress during last decades, current numerical weather prediction models are still not able to forecast convective storms with the spatial and temporal accuracy generally required by the end users. To improve the understanding of convective storms, several specific observational field campaigns have been performed (e.g., Wilson and Roberts 2006; Browning et al. 2007; Wulfmeyer et al. 2011). Nevertheless, a better knowledge of convective storm behavior based on long-term observations is also required to develop nowcasting rules (Wilson et al. 2010). Such a knowledge is also beneficial for the verification of regional climate models, the design of hydraulic infrastructures (e.g., dam, sewer systems), and more generally for all activities that are affected by these storms.

Numerous studies on the statistical characteristics of convective storms in different regions of the world have been performed. They differ by the type of weather data used, the proposed methodologies, and the final purpose of the study. Earlier works, which were based on weather station reports of heavy rain and thunder, suffer from bias due to the limited spatial representativity. Lightning detection network observations (Mäkelä et al. 2011) provide a good indication of the electrical activity of thunderstorms but miss convective storms when they
do not produce lightning. Satellite observations provide valuable information about the characteristics of convective clouds for very large areas. Unfortunately, large errors appear when deriving precipitation from infrared brightness temperature (Tuttle et al. 2008). A higher accuracy can be obtained using a satellite precipitation radar like in the Tropical Rainfall Measuring Mission (Liu and Zipser 2009), but the midlatitudes are only partially covered. Compared to satellite observations, the use of ground-based radar measurements can give better insight into the small-scale spatial structure of convective systems. Indeed, weather radar provides volumetric reflectivity measurements at high spatial and temporal resolution. With a few minutes between each scan, it can well capture the precipitation activity of the convective storms from growth to decay. Radar reflectivity measurements are affected by various sources of error (Doviak and Zrnić 2006). Calibration errors, sidelobe effects, shielding, attenuation, and ground clutter are the most important ones. Some of these errors tend to increase with distance from the radar. Furthermore, the atmospheric volume covered by the radar is limited by the cone of silence at short range and overshooting at long ranges. Attenuation can be particularly important in the presence of convective storms especially nearby the radar. Therefore a careful treatment of the radar data is required.

Two different approaches for the analysis of radar reflectivity data can be found in the literature. The Eulerian approach is based on the temporal analysis of the measured field at fixed points (typically each grid point), while the Lagrangian or object-oriented approach is based on the tracking of spatial features across successive images. The former approach is straightforward and is generally used to derive long-term statistics at larger time and spatial resolution using 2D radar data. For example, Weckwerth et al. (2011) developed a radar climatology of convective initiation and enhancements for the Convective and Orographically-Induced Precipitation Study (COPS) region based on a 2D radar composite. In the Netherlands, Overeem et al. (2009) derived a 10-yr radar climatology of rainfall. Many studies can also be found in the United States using the National Operational Weather Radar (NOWRAD) 2D radar composite (e.g., Schumacher and Johnson 2006; Carbone and Tuttle 2008; Lombardo and Colle 2010). The Lagrangian approach is more complex since it requires specific feature identification and tracking techniques. The use of data with high spatial and temporal resolution is particularly suited for this approach. Popular convective storm-tracking algorithms include Thunderstorm Identification, Tracking, Analysis and Nowcasting (TITAN; Dixon and Wiener 1993) mainly developed by the National Center for Atmospheric Research and Storm Cell Identification and Tracking (SCIT; Johnson et al. 1998) developed by the Storm Prediction Center. In Germany, Weusthoff and Hauf (2008) studied the life cycle of convective-shower cells using a specially adapted tracking algorithm on 5-min 2D radar composite data. A 6-yr study of storm tracks over northwestern Italy is proposed by Davini et al. (2011). A limited number of studies make specific use of volumetric data, mostly performed outside Europe. The size distribution of convective clouds is required for parameterization of its effects into general circulation models. An early comprehensive study of convective radar echoes can be found in López et al. (1984) for Florida using 2 years of C-band and S-band radar data and simple tracking techniques. Their results show the distribution of radar echo characteristics with a focus on individual cells, its temporal variability and life cycle dynamic. In a 2002–06 study for North Dakota based on Next Generation Weather Radar (NEXRAD) Level-III storm database (generated by SCIT), Mohee and Miller (2010) provide temporal statistics of convective storms and descriptive statistics of storm-track properties (e.g., duration, intensity, and mean speed). An extended 9-yr analysis based on both reflectivity and velocity measurement of a S-band radar in Quebec can be found in Bellon and Zawadzki (2003). They provide results on the location, strength, and frequency of occurrence of severe convective events. Saxen et al. (2008) provide a 9-yr analysis of TITAN storm tracks (duration, size, speed, intensity, and echo top) for New Mexico. The same methodology was used in Australia with additional wind data by Potts et al. (2000) and with a focus on cell height by May and Ballinger (2007).

In this study, we analyze the convective storm frequency and characteristics in and around Belgium using high-resolution volumetric C-band radar data over an observation period of 10 years. To our knowledge, such a long volumetric dataset has never been used in Europe for this kind of study. An objective methodology based on the TITAN storm tracker has been carried out to provide a robust and comprehensive analysis of those data. Both Eulerian and Lagrangian approaches are followed in the statistical analysis. The radar data, the clutter mitigation technique, and the storm-tracking system are presented in section 2. Section 3 describes the statistical methodology and the storm classification. In section 4, the spatial and temporal characteristics of the convective storm activity are analyzed (Eulerian approach). The statistical characteristics of the convective storm tracks are analyzed in section 5 (Lagrangian approach). Conclusions and perspectives are drawn in section 6.
2. Radar data and storm tracking

a. C-band radar measurements

Since 2001, the Royal Meteorological Institute of Belgium (RMIB) has operated a C-band (5.62 GHz) weather radar located in Wideumont (49.9°N, 5.5°E), in the southeast of Belgium (Fig. 1), at 592 m above sea level. The radar, which has a range of 240 km, covers Belgium and Luxembourg and also parts of France, the Netherlands, and Germany. It is a single-polarization radar with Doppler capability, used to filter ground echoes. A monitoring of the electronic calibration is performed using the mean ground clutter reflectivity at short range and the reflectivity produced by three towers in the vicinity of the radar. These point targets also allow controlling range and azimuth assignments. The radar performs a dedicated reflectivity scan at five elevations (0.3°, 0.6°, 1.8°, 3.3°, and 6.0°) every 5 min with a pulse repetition frequency of 600 Hz and pulse duration of 0.836 μs. The radar volume data have a resolution of 1° in azimuth (an average of 33 pulses) and 250 m in range (an average of two successive range bins). Every 15 min, the radar also performs a 10-elevation scan without Doppler filtering. A third scan, limited to 120 km, is used to retrieve radial velocity data. Because of the height of the radar tower and its position near the top of the Ardennes Ridge, beam-blocking effects are extremely limited. More information regarding the radar characteristics and scanning strategy can be found in Delobbe and Holleman (2006). Only the data from the first scan of the Wideumont radar are used in this study. The vertical coverage of the second scan is larger, but the repetition cycle of 15 min is too long to properly track convective storms.

The volumetric data of the Wideumont radar have been archived at RMIB since 2002. No significant changes in the radar calibration have been encountered since 2002. The parameters of the Doppler filter have been changed in 2004 resulting in a stronger clutter filtering. No postprocessing corrections have been performed on the data. The radar data availability for the most active period (from April to September) is close to 97.5%. If gaps of 15 min are tolerated, which is on the order of a storm duration, the availability reaches 98%.

FIG. 1. Elevation of the study area, with location of the Wideumont radar, country borders, and Belgian provinces.
We assume there is no convective activity for missing files in the least active period (from October to March).

**b. Clutter mitigation technique**

Weather radar measurements can be contaminated by nonmeteorological echoes such as airplanes or insects. Nevertheless, the main source of clutter is due to the radar beam (including side lobes) reaching the ground. During normal beam propagation (NP clutter), it occurs mainly at elevated places such as hills surrounding the radar. It can also appear over larger areas during particular atmospheric conditions with abnormal propagation of the beam toward the ground (AP clutter). NP and AP clutter exhibit similar radar echoes that can be interpreted as precipitation echoes. If a contiguous contaminated area is sufficiently large, it can be mistaken as a convective storm by a cell tracker system. If such an area is detected across successive radar images, it can be tracked as a convective storm. The resulting erroneous track is likely to have a very small mean speed. Since actual convective storms can be stationary, this characteristic cannot be used as a selection criterion.

The identification and mitigation of ground clutter remains a challenging problem in radar meteorology mainly because it is not always easy to distinguish precipitation echoes. A Doppler or statistical filter in the signal processor is typically used to remove near-zero velocity echoes. This kind of filter is not perfect and postprocessing of the radar data is needed for the remaining clutter. A summary of clutter mitigation techniques can be found in Hubbert et al. (2009), who recommend the fuzzy logic approach for its simplicity and practicability. This method, which is implemented in the TITAN system, uses a combination of several indicators to make a clutter elimination decision. In this study, the velocity field is not available for the first radar scan. Therefore, two estimates of the reflectivity radar field spatial variation are used as indicators. The texture feature is the mean squared difference of the reflectivity, 

$$\text{TDBZ} = \frac{1}{N} \sum_{i} (Z_i - Z_{i-1})^2,$$

where $Z_i$ is the reflectivity at range gate $i$ and $N$ is the number of gates used along the radial beam. The spatial variability of the reflectivity field (“SPIN change”) (Steiner and Smith 2002) feature measures the number of significant gradient changes along the radial beam that satisfy the conditions

$$\text{sign}(Z_i - Z_{i-1}) = -\text{sign}(Z_{i+1} - Z_i)$$

and

$$\frac{|Z_i - Z_{i-1}| + |Z_{i+1} - Z_i|}{2} > 5 \text{ dBZ}.$$  

For each indicator, an interest function is defined, which associates its value with the probability of being clutter. The interest functions used in this study are the same as in Kessinger et al. (2003). If the final indicator exceeds 0.5, the gate is considered as clutter and its value set to zero.

**c. Storm tracking**

The storm tracker TITAN (Dixon and Wiener 1993) has been developed for automatic identification, tracking, and forecasting of convective storms based on radar reflectivity measurements. TITAN is in constant development by several contributors and it evolved in a large suite of software with many capabilities. It can ingest data from various types of weather radars and other observation systems while displaying them altogether. In a first step, the radar volume data in polar coordinates are transformed using an eight-point bilinear interpolation into a 3D Cartesian grid of 0.5-km size. The storm identification algorithm defines a convective storm as a three-dimensional region with reflectivity values exceeding a given threshold. The volume of the region must be larger than a given threshold to be considered as a valid storm. The storm-tracking algorithm matches storms across two successive radar scans using combinatorial optimization. It finds the set of storm paths that minimizes a cost function, which is the sum of volume and distance weighted differences for each path. Before the optimization step, it also uses an overlapping technique to match the storms. An example of storm identification and tracking is shown in Fig. 2. The algorithm can also deal with storms that exhibit some evolution, either by merging or splitting. Merging occurs when one identified storm is matched with several storms in the previous image. Splitting occurs when several identified storms are matched with one storm in the previous image. The whole path of the storm, from genesis to decay including possible interactions, is referred to as a track and associated with a unique identification number. Tracks with or without interactions are labeled respectively as complex or simple tracks. Each individual track (i.e., component of a complex track between interactions) is associated with a second identification number. TITAN is also able to forecast storm evolution based on simple extrapolation using a linear or parabolic trend, but this feature is not the subject of this study.
The TITAN storm-tracking system is based on several parameters influencing its performance. The grid size of the three-dimensional interpolation grid has been set to 0.5 km. This is consistent with the radar resolution and provides a sufficient accuracy for storm detection. The basic and most important parameter is the reflectivity threshold used to identify a storm. The second important parameter is the volume threshold used to discard non-significant identifications. These two parameters are strongly related. A high reflectivity threshold (typically 40 dB$Z$) leads to the identification of small convective cores. This choice allows easier tracking but both growth and decay stages of the storm can be missed. A low reflectivity threshold (typically 30 dB$Z$) leads to the identification of large areas. The entire life cycle can be captured but the large storm areas may include several reflectivity peaks that are not identified individually. More generally, undesired behaviors of the tracking algorithm include dropped association, which breaks the storm track into several parts; permuted match between two storms close to each other; and wrong match between a decaying cell and a newly born cell. A performance analysis of TITAN can be found in Han et al. (2009) with some suggestions for improvements. TITAN has been examined by eye for several distinct cases. To evaluate its performance over a long period an objective method is needed. Lakshmanan and Smith (2010) propose to compute bulk statistics on the track properties that reflect the good performance of the algorithm (i.e., mean duration of the tracks, linearity of the tracks, and continuity of the storm attributes). For the Eulerian approach of this study, convective storms are identified using relatively high thresholds values of 40 dB$Z$ (minimum volume of 10 km$^3$) and 45 dB$Z$ (minimum volume of 5 km$^3$). This corresponds respectively to rain rates of 12 and 28 mm h$^{-1}$ using the default NEXRAD relationship $Z = 300R^{1.4}$, which is valid for deep convection. Those reflectivity thresholds are used in many studies that follow the same approach. The 40-dB$Z$ value is also proposed by Steiner et al. (1995) to identify convective precipitation. For the Lagrangian approach, the 40-dB$Z$ threshold is used.

3. **Statistical methodology**

a. **Storm properties**

The TITAN storm-tracking system has processed 10 years (from 2002 to 2011) of volumetric radar data.
For each identified convective storm, different characteristics are computed by TITAN. These include morphological properties such as volume, precipitation area (i.e., the area of the storm close to the ground), and top (maximum height of the 18-dBZ region). Physical quantities such as water mass or precipitation rate are derived from the reflectivity values. The recording of all the grid points that make up the identified storm was not possible because of the huge amount of data. A 2D representation of the storm envelope (i.e., the area obtained from the 3D volume projection on the surface) is used as a good alternative. A fitted ellipse, which is defined by its center, orientation, and minor/major radius is used as a first simple representation. This representation is not supposed to be realistic for a storm that exhibits a large extension in one direction. A more accurate representation is based on a convex polygon that is constructed by projecting 72 radials out of the storm envelope centroid (i.e., geometric center). Nonetheless, such polygon could fail to represent properly squall lines or bow echoes. Motion properties such as the speed and the direction are derived from successive tracked positions of the storm centroid. For the whole storm track, the mean and maximum of the storm properties are also computed. Additional properties have been computed such as the eccentricity of the fitted ellipse.

**b. Storm types**

It has been shown that severe weather produced by convective storms can be related to its morphological characteristics (e.g., Gallus et al. 2008). Various convective storm classifications have been proposed in the literature and a summary can be found in Schoen and Ashley (2011). Although storms exhibit a wide spectrum of morphologies and organizations, predominant modes of organization can be found in relation to synoptic and mesoscale forcing. Unorganized cellular convection is observed in weak shear environment while organized cellular convection (line or clusters) and supercells are associated with moderate to strong wind shear. A mesoscale convective system is defined as a contiguous convective precipitation area with a scale of more than 100 km in a given direction, which is often embedded in stratiform precipitation (Houze 2004). Regarding the huge amount of data, it is necessary to use an automated classification technique. While complex classification methods exist (e.g., neural network technique), a simple hierarchical technique based on a decision tree, is used in this study following Rigo and Llasat (2004) and Gagne et al. (2009). The classification is based on several morphological properties of the storm as shown in Fig. 3a. The size of the envelope area is used to determine the general type (i.e., top of the decision tree) of identified storms. Storms that have a 40-dBZ envelope area smaller than 100 km² tend to have a single reflectivity peak and are given the cellular type. Storms that exhibit an envelope area bigger than 100 km² are usually made of contiguous reflectivity peaks and are labeled as convective systems. Since the TITAN algorithm lacks information about the organization of convective storms, an algorithm has been developed to compute the distance between the storms. Consequently, convective cells are further classified as isolated if no neighboring storms exist at a 10-km distance. Convective cells are considered as organized in the other case. Convective systems are classified as “linear” if the eccentricity of the fitted ellipse exceeds 0.95; “nonlinear” otherwise. This classification is performed for each storm identified at a given time. It is clear that convective storm morphology and organization can evolve in time especially when splitting or merging occurs.
c. Storm selection

The performance of the storm identification and tracking algorithm is affected by the distance to the radar. At small distance, storms will not be identified because of the cone of silence, while at large distance, only storms with large vertical development will be identified. Besides, the sample volume increases at larger distance because of beam broadening. Therefore the study area is limited to ranges from 20 to 180 km and represents a total area size of about 10^5 km^2 (Fig. 1). These selection criteria will be used for the Eulerian approach (based on the identified storms). For the Lagrangian approach (based on the tracked storms) a specific track selection is needed. To ensure that the complete path of the storms is captured, storm tracks that exist after or before a missing file are discarded. Furthermore, the whole storm track is limited to ranges from 20 to 220 km. This ensures that a part of the track is not cut by the cone of silence or the border of the radar coverage. It is important to note that this spatial condition will discard long-living storms, typically mesoscale convective systems that can travel hundreds of kilometers.

4. Eulerian statistics

a. Sample description

During the whole 10-yr period, more than 1 million storms have been identified at 40 dBZ by TITAN in the study area. Taking into account the 5-min time step, it corresponds to a cumulative storm duration of 3500 days. If the 45-dBZ threshold is used, only half of these storms are identified. On Fig. 3b one can see the distribution of convective storm types. The most frequent storm type observed is the cellular type. Organized cells are 5 times more frequent than isolated cells. Convective storm systems are relatively rare (about 5%). Nonlinear systems are more frequent than linear ones.

b. Local convective storm activity

Climate variations exist within the study area because of the coastal influence, the orography, and the land cover. The probability of convective storm at a given location \( P_d(x, y) \) is calculated as the mean of the storm activity \( i_d(x, y, t) \) over a given period. At a given time \( t \), the storm activity \( i_d(x, y, t) \) is equal to 1 if \((x, y)\) is inside an identified convective storm (defined using reflectivity threshold \( d \)) and is 0 otherwise. Since the interpolated radar data have a resolution of 0.5 km \( \times \) 0.5 km, the polygonal representation of the storm is back-transformed to this regular grid. Then the probability is computed for each pixel \( (I, J) \):

\[
P_d(I, J) = \frac{1}{M} \sum_{T=1}^{M} i_d(I, J, T),
\]

where \( M \) is the total number of measurement time steps \( T \) during the period. Even though a temporal resolution of 5 min is small, a so-called jumping cell effect may occur for fast-moving storms. Its impact is supposed to be very limited since the analysis is based on a long observation period.

Figure 4 shows the probability of convective storm in the study area. For the whole 10-yr period (Fig. 4, top panel), \( P_{40} \) (\( P_{45} \)) ranges mostly between 0.04\% (0.01\%) and 0.08\% (0.03\%). The average probability is around 0.07\% (0.02\%), which corresponds to about 6 (2) h of convective activity per year. The standard deviation of the yearly \( P_{40} \) (Fig. 4, top-right panel) is 0.03\% on average, which highlights some interannual variability. This variability is higher for \( P_{45} \) (not shown) because of the smaller number of events. In the northeastern United States, Murray and Colle (2011) found a probability ranging between 0.10\% and 0.30\% for \( P_{45} \) during April–September using a 2 km \( \times \) 2 km resolution. The probability of convective storm tends to decrease with distance to the radar. This is a combined effect of signal attenuation, beam broadening, and beam overshooting. Concentric circle patterns are slightly apparent on the image and are another radar artifact related to the five-elevation scanning geometry. The probability of convective storm increases when going to the southeast of the study area. This effect seems related to the general increase in orography. Localized areas with higher probability are also observed, especially on the north and east of the radar. These areas are correlated with the orographic features shown in Fig. 1. It is unclear whether the observed spatial variations are caused by real climatic variations or by radar artifacts. The radar measurements are closer to the ground in high-orography regions, which could explain a higher detection rate of shallow convective storms. Besides, we cannot exclude that some ground echoes remain in the data even if a very careful ground clutter elimination has been performed. Ground echoes are more frequent in elevated places, which could also contribute to correlate the radar-derived probability of convective storm with the orography.

The middle panel of Fig. 4 shows the spatial variations of \( P_{40} \) for the months June, July, and August. The probability reaches 0.25\% in July and August, which corresponds to 2 h month\(^{-1}\). The relative probability tends to increase toward northwest between June and August. This might be related to the orography and the proximity of coast.
The bottom panel of Fig. 4 shows the diurnal variations of $P_{40}$ for the most active hours of the day. In the afternoon the probability increases toward east, while in the evening the probability is slightly higher in the western part of the study area.

c. Area-averaged convective storm activity

In this section we analyze the characteristics of the convective storm activity over the whole study area whose size is about $10^5$ km$^2$. Different area-averaged indices are considered: the number of identified storms $N$, their fractional coverage $C$ expressed in percent, and their total water mass $W$. If a storm lies on the border of the area, only a fraction of the storm is used to compute the indices. Here $W$ is based on an empirical relationship between the reflectivity $Z$ of the sample volume and its equivalent water mass $M$. In TITAN this relationship is $Z = 20 \times 645M^{1.75}$ for liquid water and $Z = (3.6683 \times 10^6)M^{1.416}$ for solid water.
water. The latter equation is used for reflectivity values exceeding 55 dBZ, which are very likely associated with hail. The water mass $M$ has a more physical meaning than $Z$ but it must be kept in mind that it strongly depends on the $Z$–$M$ relationship.

Figure 5 shows the probability to exceed a given index value for different periods aggregated by years, months, and hours. It is important to note that the results are shown using a logarithmic scale. Each probability is computed using all available 5-min data including dry periods. Here $N$ and $W$ are normalized to an area of $10^4$ km$^2$. Only the results for the 40-dBZ identification threshold are shown here.

The probability to observe at least one storm in the study area (which corresponds to 0.1 storm per $10^4$ km$^2$) ranges from 12% to 20% from year to year with an average of 15%. This probability is above 20% from May to August and reaches a maximum around 35% in July–August. An important daily variation is observed with a lower probability during the night and the morning followed by a steep increase from 0900 to 1500 UTC and a decrease in the evening. The diurnal maximum is the consequence of the solar heating effect that helps triggering convection during the afternoon (local solar noon is around 1145 UTC). This well-known effect has been observed in many previous studies (e.g., Murray and...
Colle 2011; Parker and Ahijevych 2007). For the sake of comparison, the same probability has been computed over a reference area of $10^4$ km$^2$, which lies within the optimal range interval of the radar (70–90 km). For this reference area, the probability is 6% for the whole study period and reaches a maximum of 14% in July–August.

The probability of exceedance decreases strongly for increasing number of storms. There is some interannual variability for the whole range of values. The term $P(N \geq 1)$, which is the probability to observe at least 1 storm per $10^3$ km$^2$, ranges from 2% in 2005 to 5% in 2006. It reaches a maximum in August with more than 10%, slightly above July and June. For situations with more than 2 storms per $10^4$ km$^2$, the maximum probability is found in June. Additionally, the diurnal variation is more pronounced and reaches a maximum at 1300 UTC. Those results are related to the effect of surface heating, which reaches a maximum around 21 June and in the afternoon. Convective storms triggered by surface heating are mainly isolated cells that can develop everywhere in the domain. This kind of situation is likely to produce the highest number of storms.

Exceedance probabilities for the fractional storm coverage $C$ exhibit a power-law behavior. For a given index value, the probability is higher during the period 2008–11 than the period 2003–05. This effect tends to increase for increasing values of $C$. The rank of year 2006, which has the highest probability for $C$ below 1%, significantly decreases for higher values. The maximum probability is found in August for $C$ below 0.2%, in June and July up to 2% and July for higher values. By definition, meso-scale convective systems (MCSs) are likely to produce the largest storm coverage. Therefore the latest result suggests that MCSs were more frequent in July during the study period. There is a relative increase of probability in the evening when $C$ increases. For $P(C > 2\%)$, the maximum probability is found between 1600 and 2000 UTC. The diurnal cycle is less clear for the highest values. Those effects might be related to the persistence of MCSs during the evening and sometimes in the night.

The distribution of $W$ is also skewed toward lower values. A similar yet higher interannual variability is found for this index. This effect is particularly important for situation with more than 1 Tg of water per $10^4$ km$^2$ with a probability much higher for the period 2008–11 than the period 2003–07. The seasonal variation is also clear for this index. An increase in the relative probability of July for increasing $W$ values is observed. This effect can also be related to the prevalence of MCSs during this month and the fact that those events are the most severe. The diurnal variation of the probability is important for the whole range of values with a slight relative increase in the evening for increasing $W$. There is a second maximum of probability around 0600 UTC for extreme $W$ values. Nevertheless, a much longer study period is required to confirm the latter result.

The interannual variability found in the analysis of the indices might be partially related to interannual variability of general circulation patterns. Allan and Zveryaev (2011) showed that summer precipitation variability over Europe is dominated by large-scale dynamics, which are associated with the summer North Atlantic Oscillation (SNAO). Especially, a positive SNAO is related to drier condition in north Europe (including Belgium) and wetter condition for south Europe. During the study period the SNAO was slightly positive on average during 2002–07 while mostly negative from 2008 onward. The SNAO was extremely variable during summer 2006, which can be related to the contrasting results obtained for that year. A similar attempt to link convective activity with remote oceanic forcing can be found in Murray and Colle (2011). It is worth pointing out that our record is too short to derive any definitive conclusion.

Some extreme situations have been recorded during the study period. At 1255 UTC 8 August 2008, almost 8 storms per $10^4$ km$^2$ have been observed in the study area. It was a typical case of isolated convective cells triggered by solar heating in the afternoon. During 14 July 2010 at 1755 UTC, the storm coverage exceeded 8% of the study area. That day, a quasi-linear convective system caused strong winds and tornadoes (Hamid 2012). The maximum total water mass was recorded during 28 June 2011 around 1700 UTC. Several meso-scale convective systems had developed in a very high CAPE environment. That day severe wind, heavy rain, large hail, and a tornado were reported in the study area. It is interesting to note that those extreme events occurred during the 2008–11 period when convective activity indices were higher on average.

The exceedance probability grouped by year, month, or day does not characterize entirely the convective activity. It is also important to investigate its internal variability at all time scales. Convective storm activity is characterized by the succession of active and inactive periods. To analyze this intermittent behavior, a convective episode is defined as the interval of consecutive time steps with nonzero indices. It is worth pointing out that this definition strongly depends on the study area since convective storms can move out of it. Figure 6a shows the empirical cumulative distribution function of convective episode duration for each year. The distribution tends to be positively skewed with about 70% of the episodes lasting less than 1 h while only about 5%
exceed 5 h. The longest episode inside the study area has been recorded during June 2008 and lasted 56 h.

For the number of storms $N$, the autocorrelation of the time series is showed in Fig. 6b. It measures the correlation of the index for a given time lag between two values of the time series. To ensure the stationarity of the time series, the analysis is performed for each year separately and focused on April–September months. It is shown that the correlation decreases strongly between 1 and 6 h confirming that most of the convective episodes are short. The correlation is higher for 24 h lag than 12 h lag because of the diurnal cycle. On average, there is a slightly higher correlation for lag of 1 day than lag of 2 days or more. This implies there is a slightly higher probability of convective storm activity on a given day when the day before was active.

To better analyze the variability of the convective activity for all time scales involved, a spectral decomposition has been performed as seen in Fig. 6c. The spectrum exhibits a power-law behavior that is represented by a straight line on a logarithmic scale plot. This is a characteristic of a scale-invariant process, which means the variability of the process at different time scales are related. Different scaling regimes (i.e., time-scale ranges) can be observed. From lower frequency ($T = 6$ months) to intermediate frequency ($T = 1$ day), a relatively flat spectrum can be seen, which can be associated with the intraseasonal variability. A second scaling regime, ranging from intermediate to higher frequencies ($T = 1$ h), can be associated with frontal disturbances. The scaling break around the 1-day period may be due to the finite size of the study area. Those results are consistent with previous studies of precipitation time series (e.g., Verrier et al. 2011). The peak at the 1-day period is associated with the diurnal cycle.

For the sake of comparison, the temporal resolution of convective index time series has been reduced to 1 day using the mean (the results are not shown on the figures). A day is considered a convective storm day if a cumulative storm duration of one hour is observed. For the part of the study area corresponding to Belgium, the yearly probability of convective storm days ranges from 24% to 33% and is at a maximum in July and August around 56%. From thunder reports in Belgium during 1971–88, the probability of thunderstorm days is 22% with a maximum at 38% in May and June. The lower figures obtained with the stations are related to their limited spatial coverage and the fact that the reports are based on lightning only. From 11 synoptic stations in southwest Germany with coverage of about $10^4$ km², Kunz et al. (2009) found a probability of thunderstorm days ranging between 25.7% and 44.3% during April–September.

d. Storm properties

In this section the distribution of identified storms characteristics are analyzed. Since no large geographical variations of storm properties have been found (not shown) the analysis is based on all storms identified in the study area. Figure 7 shows the interannual, seasonal, and diurnal variations of the time-averaged number of storms that exceed a given volume, echo top, or mass. This number, which is computed for the whole study area, is normalized to an area of $10^4$ km².

The distribution of storm volume follows a power law, which implies that smaller storms are more often seen than bigger storms. The distribution is relatively similar from year to year for smaller storms. However, the average number of storms bigger than $1000$ km$^3$ is about 4 times as high for the years 2008–11 as for the
years 2004 and 2005. The average number of storms exceeding a given volume is higher in June and July. This effect is much more marked for bigger storms. For volume increasing to about 2000 km$^3$, a constant difference of average number is seen between the night minimum and the maximum. This maximum occurs 1 h later for biggest volumes than for small volumes. In the evening, there is an increase in the relative average number of storms for increasing volume. This result might also be related to the persistence of MCSs, which have bigger volumes by definition.

The storm echo top is a good indication of the storm severity. It is calculated using the 18-dBZ reflectivity threshold. Obviously, the actual top of the cloud can be higher. Most of the identified storms have an echo top below 5 km. In July, the mean number of storms per 10$^4$ km$^2$ area whose echo top exceed 10 km is 0.01. The interannual variability of the probability slightly increases with increasing values. For echo top above 5 km, the average number of storms is higher in June–July than in May and August. Contrary to the volume, the diurnal variation becomes larger with increasing tops and no increase in the evening is observed.
The storm mass corresponds to the amount of water content inside the storm. The distribution of this characteristic also follows a power law with the most intense storms being relatively rare. The interannual variability increases with increasing mass. For a given mass of 2 Tg, the mean number of storms is above 0.01 for 2008–11 whereas it is close to 0.005 for 2004 and 2005. For exceedance values above 0.2 Tg, the mean number of storms is 2 times as high in June–July as in May and August. However, for storms exceeding 10 Tg, it is also maximum in May. The mean number of storms is highest during the afternoon for all values. Nevertheless, extreme values are observed at any time of the day.

The power-law behavior of storm properties found in this study is consistent with the results of previous studies (e.g., May and Ballinger 2007; Saxen et al. 2008) and can be seen as a general characteristic of convective storms. The results obtained for higher values of storm characteristics should be taken with care since the sample of storms becomes relatively small.

5. Lagrangian statistics

a. Sample description

About 3 million storm tracks have been recorded in the extended study area during the 10-yr period. The percentage of those tracks impacted by missing files is 2%. A few of them (10%) reach the boundary of the study area. About 60% of the tracks last less than 15 min and are discarded. After applying all selection criteria, the number of storm tracks is about 1 million, including 80% of simple tracks. The 20% of complex tracks (i.e., with merging or splitting) are mainly made of a few individuals tracks (Fig. 8a).

b. Storm-track kinematics

In this section we analyze the kinematic of storm tracks. The storm tracks are divided between simple and complex tracks. Furthermore, individual tracks within complex tracks are also analyzed. Individual tracks that last less than 15 min are discarded. On Fig. 8b one can see the empirical cumulative distribution of storm-track duration. The mean storm duration is 25 min for simple tracks and 70 min for complex tracks. Simple storm tracks are mostly short lived with only 5% lasting more than 50 min. Complex storm tracks, which are made of several individual tracks that interact, last more than 100 min with a probability of 20%. Only 3% of complex tracks exceed 200 min. It is interesting to note that the distribution of duration is similar for individual tracks and simple tracks. The tail of the distribution might be influenced by the limited size of the study area.

Figure 8c represents the distribution of the storm mean direction for both simple and complex tracks. It is computed using the successive positions of the storm volume centroid. It shows the prevalence of the northeast direction. This result is consistent with the distribution of surface wind direction prevailing during convective situations. A perfect match is not expected since the steering wind at given altitude can differ significantly from the surface wind.

The distribution of storm tracks’ mean speed (Fig. 9a) follows a Weibull distribution. This is also the typical distribution of surface wind speed measurements. About half of the simple storm tracks exhibit a mean speed below 30 km h⁻¹. Their probability to exceed 60 km h⁻¹ is only 5%. The mean speed of individual tracks and complex tracks tends to be slightly higher. The analysis of the seasonal (Fig. 9b) and diurnal (Fig. 9c)
variations shows that their motion is slower in the summer and in the afternoon.

c. Storm-track initiation

We analyze here the spatial variations of convective storm initiation (e.g., Wulfmeyer et al. 2011) using both 40- and 45-dBZ thresholds. In TITAN storm initiation corresponds to the position of the storm at the beginning of the track. The statistical framework of spatial point process (Diggle 2003) is used. The storm volume centroid is used as a point representation of the 3D storm shape. This approximation is justified for convective cells but is less valid for bigger convective systems. A strong correlation between storm initiation locations is expected at short time scales given the predominance of organized convective storms. This effect is supposed to be very limited over a relatively long dataset. Figure 10 shows the density of convective storm initiation for the whole study period. It is obtained by using a kernel smoothing on the spatial point pattern. The analysis is performed using individual tracks (40 and 45 dBZ) and simple/complex tracks (45 dBZ). All tracks starting in the study area are taken into account. The three figures exhibit similar characteristics. Localized areas with higher or lower density are observed and correspond to orographic variations. The storm initiation spatial point pattern can be further analyzed using a test for complete spatial randomness (CSR). Our method is based on the distribution of nearest-neighbor distance (not shown), which measures the dispersion or clustering of a spatial point pattern. This empirical distribution is compared to the Poisson distribution using the Clark–Evans statistical test. To mitigate the effect of the distance, the study area has been divided into range rings of 20-km width. For all rings, the results of the test of CSR are negative and the alternative hypothesis of clustering is retained. This suggests that there exist preferred areas for initiation. Note, however, that the distribution is still very

![Fig. 9](image_url)

Fig. 9. (a) Estimate density of storm-track mean speed; time-averaged number of identified storms exceeding a given speed grouped by month and by hour.

![Fig. 10](image_url)

Fig. 10. Density of storm initiation: individual tracks (including simple tracks) at (left) 40 dBZ and (center) 45 dBZ; (right) simple and complex tracks at 45 dBZ.
close to the Poisson distribution, meaning that the clustering is limited.

6. Conclusions and perspectives

The characteristics of the convective storm activity in and around Belgium have been analyzed over a 10-yr period (2002–11). The analysis is based on high-resolution volumetric radar observations that have been filtered by a fuzzy logic clutter mitigation technique. The TITAN algorithm has been used to identify and track convective storms on successive 5-min radar scans using a reflectivity threshold of 40 dBZ. In a study area limited to between 20 and 180 km from the radar, more than 1 million convective storms have been identified, which represents a cumulative duration of 3500 days. A simple classification has been proposed that reveals that most storms are of the cellular type, while convective systems are relatively rare. Convective cells tend to be organized rather than isolated.

The average probability of convective storm activity at a given location is 0.07%, which corresponds to 6 h yr$^{-1}$. Besides radar artifacts, the spatial variations of the probability are relatively limited. However, some regions with slightly higher or lower activity can be identified and are related to the effect of orography. There is an interannual variability with an averaged standard deviation of 0.03%. The relative probability trends to increase in the western part of the study area from May to August and from the afternoon to the evening.

The probability to observe at least one storm is 15% for the whole study area whose size is 10$^5$ km$^2$. A 6% probability is obtained for a 10$^4$-km$^2$ reference area. The probability reaches a maximum in July and August, with respectively 36% and 14% for the study area and the reference area. Most of the time during those active periods, the number of storms, their fractional coverage, and their total water mass remain limited. The probability that one of these characteristics exceeds a certain level is higher for the year 2002 and the period 2008–11 than for the period 2003–07. This effect is more pronounced for severe events especially for the storm mass. It might be related to variations in general circulation patterns, which are influenced by oceanic forcing such as the NAO. During the year, a high number of storms and large storm coverage are found, especially in June when surface heating reaches a maximum. Events with very large storm coverage and with large total water mass are found mainly in July when most MCSs have been recorded. The probability to observe a high number of storms at the same time is highest in the early afternoon around 1300 UTC. The probability that storms cover large areas and have a huge mass of water is higher in the late afternoon and also in the evening. Because of the persistence of MCSs, the most severe situations in terms of total area and water mass can occur at any time of the day.

The occurrence of the activity in the study area is highly intermittent and its autocorrelation decreases quickly from 1 to 6 h. There is a slightly higher probability of convective storm activity on a given day when the day before was active. The frequency analysis of the time series reveals a scaling behavior with two distinct regimes, which can be related to the seasonal and synoptic variability.

No important variations in the storm characteristics have been found across the study area. Most of the identified convective storms have a relatively small volume, with a limited echo top, and they generate a small amount of water. On average storms tend to be bigger, deeper, and more intense in June and July. The diurnal variation increases when the echo top increases while it is shifted toward the evening for increasing volume and water mass. The average number of severe storms does not vary substantially from May to August and during the day.

Storms tend to be short-lived with a mean duration of 25 min for simple tracks and 70 min for complex tracks. Individual tracks that are part of complex tracks follow the same distribution as simple tracks. Convective storms move preferably toward a northeast direction with a median speed of 30 km h$^{-1}$ and tend to be slower in summer and in the afternoon. Regions with slightly more or slightly less storm initiations have been found and are related to orographic variations.

Despite the limitations inherent to radar observations, this analysis gives an unprecedented view of the convective storm characteristics in the region covered by the radar. These characteristics are probably representative of a larger region, including the north of France and Germany where the same methodology could be applied. This study, which was mainly observational, offers several perspectives for better understanding of the convective storm activity. First, additional observations from neighboring radars should be included in the analysis. In overlapping areas, the quality of the observations will be improved. Furthermore, this will allow the discrimination between radar artifacts and true meteorological effects. It would be interesting to analyze the results obtained using other tracking methods (e.g., SCIT). The Lagrangian approach can be extended by studying the evolution of storm characteristics and their correlation. Additional storm properties could be analyzed by including other observations like lightning activity, cloud characteristics derived from satellite observations, and water vapor fields derived from GPS.
Also, the relation between convective storm activity and weather regimes could be investigated.

**Acknowledgments.** The authors thank the reviewers for their very constructive comments and suggestions that allowed a substantial improvement of the paper. We thank also Mike Dixon (UCAR, United States) for his support concerning the use of TITAN. The comments of Maryna Lukach, Dieter Poelman, and Maarten Reyniers were highly appreciated. This research has been supported by the Belgian Federal Science Policy (Research project ACTION 1 MO/34/018).

**REFERENCES**


