BEEhaviourLab: A high-throughput platform for sublethal stressor screening in insects

Rachel H. Parkinson[1,2,*], Oliver N. F. King[3], Jung Chun (Zaza) Kuo[4], Kieran Walter[1], Ash Silva[1], Jennifer Scott[1], Cait Newport[1], Geraldine A. Wright[1], and Stephen Roberts[5].

Affiliations

1. Department of Biology, University of Oxford

2. School of Biological and Behavioural Sciences, Queen Mary University of London

3. OxRSE, Doctoral Training Centre, University of Oxford

4. Royal (Dick) School of Veterinary Studies, University of Edinburgh

5. Machine Learning Research Group, Department of Engineering Science, University of Oxford

*: corresponding author

R.H. Parkinson: rachel.parkinson@biology.ox.ac.uk; r.parkinson@qmul.ac.uk

Abstract

Chemical risk assessment for insects relies largely on mortality endpoints in a few model species, limiting detection of ecologically relevant sublethal effects and cross-taxon comparisons. Behaviour is a sensitive indicator of neurotoxic stress, but scalable, standardised measurement has remained challenging. We present BEEhaviourLab, a low-cost, automated platform for high-throughput behavioural phenotyping. The system integrates parallel video and audio recording with computer vision tracking, acoustic classification, experimental control, and automated analyses, enabling long-duration experiments across many replicates and species. A single lightweight object-detection model tracks multiple untagged insects simultaneously, including individuals from different taxa (e.g., bees and hoverflies). Using this platform, we quantify species-specific circadian activity patterns and assess toxicity of the widely used veterinary pesticide moxidectin in the bumble bee (*Bombus terrestris*). Acute contact exposure caused dose-dependent reductions in locomotion and buzzing at sublethal concentrations, with acoustic measures more sensitive than video. These results demonstrate that scalable, multimodal behavioural phenotyping can detect biologically meaningful neurotoxic disruption well below lethal thresholds, providing a practical path to integrate sublethal endpoints into chemical risk assessment.

Keywords

Sublethal toxicity; Behavioural phenotyping; Automated insect tracking; Pollinators; High-throughput screening

1. Introduction

Chemical regulation relies on standardized toxicity testing to assess risks to human health and the environment. For terrestrial invertebrates, environmental risk assessment is typically based on acute oral or contact assays conducted in a limited set of model species, most commonly the honey bee [1]. While lethal dose estimates derived from these assays are reproducible and comparable across studies, they capture only a fraction of the biological effects induced by chemical exposure [2]. Sublethal impacts – particularly those affecting behaviour, development, immunity, and communication – are increasingly recognised as ecologically relevant and capable of propagating to population- and ecosystem-level outcomes [3]. However, such effects remain difficult to quantify in a standardised, scalable, and regulatorily defensible manner, limiting their integration into routine risk assessment [2]. This gap is particularly concerning for insect pollinators. Bees underpin biodiversity, food security, and ecosystem resilience [4,5], yet they are exposed to diverse chemical stressors, including pesticides applied to crops and veterinary medicines used in livestock [6,7]. Many of these compounds persist in soils, water, nectar, and pollen at biologically relevant concentrations [8,9]. Veterinary endectocides are of particular interest in pasture systems, where residues excreted in dung can create sustained exposure routes for ground- and cavity-nesting species.

Moxidectin is a widely used macrocyclic lactone for parasite control in livestock, companion animals, and humans (Schraven et al. 2021). It acts primarily on glutamate- and γ-aminobutyric acid (GABA)-gated chloride channels in invertebrates, increasing chloride conductance, hyperpolarising neurons, and suppressing synaptic transmission [10,11]. Because these channels regulate sensory processing and motor control, macrocyclic lactones are expected to disrupt behaviour at sublethal doses. Following livestock treatment, moxidectin is excreted largely unchanged and can persist in dung at hundreds of ng/g concentrations for many weeks [12].

Consistent with its mode of action, moxidectin impairs locomotion in non-target insects: in adult dung beetles it induces dose-dependent ataxia and paralysis, alongside disruption of olfactory responses [13]. More broadly, compounds in the macrocyclic lactone class show high acute toxicity to bees, with topical LD_{50} values in the low tens of ng per bee reported for honey bees (*Apis mellifera*) and blue orchard bees (*Osmia lignaria*) [14,15]. Field studies further demonstrate that moxidectin accumulates in cavity-nesting Hymenoptera (including leafcutter bees and mason wasps), with higher concentrations detected in nonviable larvae than viable larvae, indicating environmentally realistic exposure capable of affecting developing brood [16]. Despite this evidence, and the widespread use of moxidectin, its sublethal effects have not been experimentally characterised in any bee species, including ground-nesting bees like bumble bees that are likely to encounter soil residues.

More broadly, ecotoxicology lacks a quantitative framework that links sublethal behavioural impairments to lethal endpoints across exposure gradients and enables robust comparison of toxic effects across species beyond standard model organisms [17]. Among behavioural endpoints, locomotion provides a particularly informative and mechanistically interpretable endpoint for neurotoxic compounds [18–21]. Even modest reductions in movement can scale into

substantial fitness consequences by limiting foraging efficiency, navigation, and predator avoidance.

Over the past two decades, advances in computer vision enabled automated behavioural and toxicological phenotyping in several model systems, including the nematode worm *Caenorhabditis elegans* [22,23], zebrafish (*Danio rerio*) [24,25]. and aquatic invertebrates such as *Daphnia* [26]. In bees, recent approaches have combined in-hive behavioural tracking with semi-field foraging assays, providing rich datasets spanning individual and colony-level processes [27]. Other laboratory-based systems, such as BumbleBox, enable tracking of individually tagged bumble bees within microcolonies and are well suited to studying nest behaviour over extended periods [28]. In *Drosophila*, "ethoscopes" enable closed-loop behavioural tracking at scale [29], and have been successfully applied to pharmacological and toxicological screening, where simple locomotor metrics can outperform more complex behavioural descriptors in predicting pesticide class [30]. Automated larval tracking platforms such as LarvaTagger have enabled the screening of hundreds to thousands of compounds with behavioural classification at scale [31]. However, the field still lacks a dedicated platform for standardised, high-throughput behavioural toxicology across insect pollinator species.

Here, we introduce BEEhaviourLab, a low-cost, scalable, automated platform for high-throughput behavioural phenotyping in insects. The system enables long-duration, standardised recording across multiple species and replicates under controlled exposure conditions. Integrated analysis pipelines generate quantitative behavioural endpoints suitable for comparative toxicological screening. We demonstrate its cross-species flexibility and its sensitivity for detecting sublethal behavioural effects in bumble bees exposed to moxidectin. By bridging the gap between mortality-based assays and complex semi-field experiments, BEEhaviourLab provides a scalable and regulatorily relevant framework for incorporating behavioural endpoints into chemical risk assessment.

2. Results

2.1. A single multi-species detector enables accurate, tag-free tracking

We developed BEEhaviourLab for automated, high-throughput behavioural tracking of insects using video and audio recordings (Figure 1). Insect tracking was performed using a computer vision algorithm, YOLO version 8 "nano" (YOLOv8n, github.com/ultralytics/ultralytics), to detect insects without the use of tags. To test generalisation across taxa, we selected four ecologically relevant and experimentally tractable insect pollinators spanning gradients in body size (10.4–15.6 mm), morphology, social organisation (eusocial and solitary), and flight dynamics: *Bombus terrestris* (eusocial bumble bee; mean length 15.6 mm), *Apis mellifera* (eusocial honey bee; 13.3 mm), *Colletes hederae* (solitary bee; 12.2 mm), and *Eristalis tenax* (solitary hoverfly with a distinct dipteran body plan and flight kinematics; 10.4 mm).

A YOLOv8n model trained exclusively on *B. terrestris* images (model 1, 155 annotated images) achieved high detection performance for bumble bees (precision, recall, and F1 > 0.97; Table 1) but generalised poorly to other taxa. Performance was lowest for *C. hederae* (precision and recall

< 0.50) and was moderately reduced for *A. mellifera* and *E. tenax*, particularly under motion blur. We therefore trained a second model (model 2) using 455 annotated images from all four species (bumble bee: 155 images; other species: 100 images). This multi-species model achieved consistently high performance across taxa (mean average precision > 0.98 for all species; Table 1), demonstrating robust detection across differences in body size, morphology, and movement dynamics. Together, these results show that incorporating morphological diversity during training enables a single lightweight detector to support accurate, tag-free tracking across pollinator taxa within a shared experimental framework.

2.2. Comparative tracking reveals species-specific circadian activity profiles

Using the multi-species detection model, we quantified locomotor activity over a 48 h period for species recorded in the BEEhaviourLab system. All species showed diel modulation, with reduced activity at night and increased activity during the day (Figure 2 A). However, species differed substantially in both their overall activity levels and the strength of the day-night modulation.

Species varied in the probability of movement as well as in locomotion speed when active. Bumble bees exhibited the highest probability of movement (speed exceeding the empirically defined noise threshold, P = 0.76, 95% CI 0.68–0.82), whereas honey bees showed the highest locomotion speeds among active individuals (2.16 mm s^{-1}, 95% CI 1.92–2.43, confidence intervals were derived from generalized linear mixed models using Wald approximations). Hoverflies were both less likely to move and moved more slowly when active. Across species, activity declined modestly over time, but this effect was small relative to interspecific differences.

Beyond differences in mean activity, locomotion exhibited strong circadian rhythmicity. A harmonic mixed-effects model revealed highly significant 24-h periodic structure, with pronounced species-specific differences in circadian amplitude (sin24: $\chi^2 = 58.0$, $p = 1.6 \times 10^{-12}$) and waveform (cos24: $\chi^2 = 162.1$, $p < 2 \times 10^{-16}$). Honey bees and ivy bees exhibited strong, high-amplitude circadian rhythms consistent with predominantly diurnal activity, whereas bumble bees maintained elevated activity across the diel cycle, resulting in a more attenuated rhythm. Hoverflies displayed comparatively weak and irregular circadian structure.

Buzzing behaviour was assessed by automatically classifying audio frames (0.1 s windows) as containing buzzing or not, using an energy-based threshold (see Methods), and reporting the proportion of time with buzzing per 60 s recording. We found species-specific differences in buzzing behaviour over time (Figure 2 D). Bumble bees exhibited the highest mean activity among non-zero observations (Figure 2 E, 0.208, 95% CI 0.163–0.261), significantly exceeding ivy bees (0.102, 95% CI 0.063–0.161; Holm-adjusted p = 0.039, EMM) and marginally exceeding hoverflies (0.111, 95% CI 0.070–0.173; p = 0.058, EMM). Honey bees showed intermediate activity (0.141, 95% CI 0.117–0.169) and did not differ significantly from the other species after correction.

In the zero-inflation component (zero buzzing activity over an entire recording, Figure 2 F), species also differed in the probability of complete inactivity. Bumble bees had a lower probability of structural zeros (0.067, 95% CI 0.052–0.086) than honey bees (0.127, 95% CI 0.112–0.144; p

< 0.001) and ivy bees (0.224, 95% CI 0.178–0.278; $p < 0.001$). Structural zeros were effectively absent in hoverflies, leading to unstable zero-inflation estimates. Across species, acoustic activity increased modestly over time ($\beta = 0.071 \pm 0.018$ SE, $p < 0.001$).

2.3. Dose-dependent locomotor impairment is driven by reduced movement probability

We next examined the effects of acute moxidectin exposure on bumble bee locomotion (Figure 3 A). Following contact exposure to moxidectin (DMSO with 0.782, 7.82, 78.2, 782 ng bee^{-1} moxidectin) or a vehicle control (DMSO only), we recorded the activity of bumble bees (*Bombus terrestris*, n=3 bees per cage, 10-12 cages per treatment) using BEEhaviourLab.

We used a two-part mixed-effects framework that separates the probability of movement from movement speed conditional on activity. Increasing dose caused a strong, monotonic reduction in the probability of movement (log-dose effect: $\beta = -0.91 \pm 0.12$ SE, $z = -7.58$, $p < 10^{-13}$), with movement probability declining from 0.66 in controls to 0.50 at the lowest tested dose and to 0.13 at the highest dose (Figure 3 B). Among bees that were active, locomotion speed also declined with increasing dose ($\beta = -0.13 \pm 0.03$ SE, $z = -4.49$, $p = 7 \times 10^{-6}$), but the magnitude of the effect was smaller (Figure 3 C). Increasing moxidectin dose also altered the circadian structure of activity. Interactions between dose and both harmonic components were highly significant (sin24: $\chi^2 = 56.3$, $p = 1.8 \times 10^{-11}$; cos24: $\chi^2 = 36.1$, $p = 2.8 \times 10^{-7}$). While the lowest dose had no detectable effect on circadian organisation, higher doses increasingly distorted the amplitude and phase of the activity cycle, with the strongest disruption observed at the highest dose (782 ng bee^{-1}).

In contrast, moxidectin had no detectable effect on inter-bee distance, a proxy for social spacing. Neither dose nor time influenced inter-individual spacing, and confidence intervals indicated that any undetected effects were small (dose: $\beta = -0.011 \pm 0.027$ SE, $z = -0.42$, $p = 0.68$; time: $\beta = 0.001 \pm 0.004$ SE, $z = 0.26$, $p = 0.79$).

2.4. Acoustic analysis reveals dose-dependent behavioural impairment

Increasing moxidectin dose revealed decreased buzzing (Figure 3 D). Dose had a strong positive effect on the probability of complete inactivity (Figure 3 E, $\beta = 0.514 \pm 0.022$ SE, $z = 23.7$, $p < 2 \times 10^{-16}$), resulting in an approximately linear increase in inactivity on the log-dose scale. Even the lowest tested dose increased the proportion of zero-activity observations by approximately 8% relative to controls, with progressively larger increases at higher doses. Among non-zero observations, buzzing activity declined more modestly with increasing dose and over time (Figure 3 F, $\beta = -0.232 \pm 0.057$ SE, $z = -4.10$, $p = 4.1 \times 10^{-5}$) and over time ($\beta = -0.0495 \pm 0.0144$ SE, $z = -3.43$, $p = 6.0 \times 10^{-4}$). These results indicate that moxidectin primarily affects buzzing behaviour by increasing the likelihood that bees do not buzz at all, rather than by uniformly reducing buzzing persistence once buzzing has begun.

2.5. Integrating locomotion and buzzing yields a sensitive sublethal toxicity threshold

To integrate information across behavioural modalities, we defined a sublethal toxicity index combining dose-specific deviations from control in movement probability, locomotion speed, and buzzing probability. Components were scaled by natural between-individual variability observed in control bees. The resulting sublethal toxicity index increased monotonically with dose, indicating progressively larger behavioural disruption across modalities (Figure 4 A). We defined the lowest observed effect concentration (LOEC) as the lowest dose at which the multimodal toxicity index exceeded 1 (i.e., the mean predicted effect across components was at least one control SD away from baseline). An index ≥1 therefore reflects a behavioural deviation exceeding natural baseline variability, establishing a standardized effect-size threshold for comparison with lethal responses. The index crossed this threshold at an estimated dose of 2.73 ng bee^{-1} (log-linear interpolation between tested doses). At higher doses, behavioural disruption exceeded the range observed in control individuals. Contributions to the composite index were not uniform across components (Figure 4 B). Reductions in movement probability contributed most strongly across doses, followed by reductions in buzzing probability, whereas changes in locomotion speed among active bees contributed least.

2.6. Behavioural disruption precedes feeding suppression and mortality

Finally, we quantified the effects of moxidectin on survival and feeding behaviour over the 48 h period. Moxidectin caused a clear, dose-dependent increase in mortality, with a 48 h LD_{50} of 81.2 ng bee^{-1} (Figure 4 C). Complete mortality occurred at the highest dose (7820 ng/bee) within 1–12 h – this dose was excluded from the behavioural assay. Feeding behaviour, measured as total sucrose consumption per cage, also declined with increasing dose. Reduced consumption relative to controls was detectable from 78.2 ng bee^{-1} onward (Figure 4 D). Feeding suppression and mortality occurred at concentrations over an order of magnitude higher than those at which clear behavioural disruption was detected. This separation highlights the sensitivity of behavioural endpoints for identifying sublethal neurotoxic effects below conventional lethal thresholds.

3. Discussion

This study introduces BEEhaviourLab as a scalable, low-cost platform for automated behavioural tracking in insects, and demonstrates its utility for both comparative behavioural analysis and sensitive detection of sublethal toxicological effects. By combining distributed hardware with automated video and audio analysis, the system bridges a persistent gap between high-throughput screening and ecologically relevant behavioural endpoints.

Recent work has established that automated behavioural monitoring can reveal pesticide-induced alterations in locomotion, circadian organisation, and colony function at doses below lethal thresholds [27,30,32]. Our results are consistent with this emerging literature, reinforcing the sensitivity of behavioural endpoints relative to mortality-based assays. However, BEEhaviourLab extends these approaches in two important ways: first, by integrating video and audio modalities within a unified analytical framework that separates behavioural state transitions from conditional

performance; and second, by enabling standardised, cross-species behavioural phenotyping using a single multi-species detection pipeline.

3.1 Scalable, standardised behavioural tracking enables comparative insect phenotyping

A key strength of BEEhaviourLab is its ability to generate long-duration, high-resolution behavioural time series at scale using inexpensive and readily available hardware. Individual recording units can be deployed and operated in parallel, enabling large numbers of replicates to be recorded under tightly standardised conditions. This substantially lowers the logistical and financial barriers to behavioural experiments that would otherwise be constrained by labour or infrastructure.

Crucially, the system performs robustly across multiple insect taxa with different body sizes, morphologies, and movement patterns. Training a single lightweight object-detection model on a modest (~100 training images per species) but diverse image set enabled accurate tracking across species without physical tagging or species-specific hardware modifications. By standardising both data acquisition and analysis, BEEhaviourLab supports direct, quantitative comparison of behavioural traits across taxa.

The species-level locomotion analysis illustrates the value of this approach. Differences in locomotor behaviour were not captured by a single axis of activity, but instead reflected distinct behavioural strategies combining movement probability and movement speed. Bumble bees were active more frequently but moved at moderate speeds, whereas honey bees showed lower activity but higher speeds when active. Hoverflies exhibited both low activity and low speed. This dissociation highlights the importance of separating behavioural state (active vs inactive) from performance once active, as mean speed alone would obscure biologically meaningful differences. The two-part modelling framework therefore provides a general and sensitive approach for comparative behavioural analysis in automated tracking datasets.

Together, these features position BEEhaviourLab as a practical tool for systematic behavioural screening in contexts where throughput, reproducibility, and cross-study comparability are limiting factors.

3.2 Sublethal effects of moxidectin reveal risks of veterinary pesticides to pollinators

Using BEEhaviourLab, we identified pronounced sublethal behavioural effects of moxidectin. Acute contact exposure to moxidectin caused strong, dose-dependent reductions in bumble bee locomotor activity at concentrations within the range reported in field residues [12]. These impairments persisted across the 48 h recording period, indicating sustained neurobehavioural dysfunction. The marked separation between behavioural disruption and the 48 h LD_{50} highlights a limitation of mortality-based assessments for neuroactive veterinary compounds.

Mechanistically, the locomotion data support a state-dependent mode of action in which moxidectin primarily increases time spent inactive, with a secondary reduction in movement speed among bees that remain active. This distinction is important: bimodal speed distributions

can obscure low-dose effects in single-distribution analyses, where changes in state occupancy dominate the mean. By explicitly separating movement probability from conditional speed, the analysis resolves low-dose behavioural suppression that aligns with the audio-derived endpoints and provides a clearer interpretation of sublethal neurotoxic effects.

Acoustic analysis converged with locomotor effects. Moxidectin increased the probability of complete buzzing inactivity and reduced buzzing persistence among active bees, again demonstrating that binary behavioural state transitions are particularly sensitive to low-dose exposure. The concordance of locomotor and acoustic endpoints strengthens inference that the observed impairments reflect genuine neurobehavioural disruption rather than artefact of a single measurement modality.

Circadian organisation was also progressively distorted with increasing dose. While control and low-dose bees exhibited robust 24 h rhythmicity, higher doses reduced amplitude and altered waveform shape, suggesting disruption of neural processes governing rhythmic behavioural regulation rather than simple global suppression. Similar circadian attenuation has been reported for other pesticides, such as neonicotinoids [33], reinforcing the sensitivity of rhythmic structure as an indicator of sublethal neural perturbation.

To integrate these effects, we constructed a multimodal sublethal toxicity index combining cycle-averaged deviations in movement probability, locomotion speed, and buzzing probability, each scaled by natural between-individual variability in controls. The index exceeded the predefined abnormality threshold at an estimated dose of 2.73 ng bee^{-1}, compared with a 48 h LD_{50} of 81.2 ng bee^{-1}. This >30-fold separation demonstrates that multimodal behavioural phenotyping can detect functionally meaningful impairment at concentrations far below conventional lethal endpoints and provides a quantitative bridge between sublethal behavioural disruption and standard toxicity benchmarks.

3.3. Limitations

Several limitations should be considered when interpreting these results. First, tag-free insect tracking does not maintain individual insect identity between recordings, and trajectories were linked using nearest-neighbour continuity rather than an explicit motion model. Identity swaps may therefore occur during close interactions, and gaps in detection were handled conservatively by carrying positions forward rather than by interpolating, which could introduce bias into downstream motion metrics.

Second, species-level circadian comparisons were confounded by differences in rearing history: bumble bees were laboratory-reared, whereas other species were wild-caught. Laboratory rearing may alter circadian entrainment or behavioural flexibility, limiting the extent to which circadian differences can be attributed to intrinsic species traits in the present study.

Third, exposures were acute and conducted under simplified laboratory conditions. Although this design enables precise dose–response estimation, it does not capture chronic exposure

dynamics, social feedback, or task allocation present in field or colony contexts. Behavioural disruption detected here therefore represents only one dimension of ecological impairment.

Finally, behavioural analyses focused on a limited set of endpoints. While these were chosen for robustness and cross-modality integration, richer behavioural features such as posture changes and fine-scale motor or acoustic signatures remain to be explored.

3.4. Conclusions and outlook

This study demonstrates that automated, high-throughput behavioural phenotyping can detect mechanistically coherent sublethal effects at concentrations below those identified by mortality-based assays. By combining multimodal behavioural measurement with effect-size standardisation, BEEhaviourLab provides a quantitative bridge between behavioural impairment and conventional toxicity benchmarks.

The framework is readily extensible. Continuous trajectory and audio data support behaviour segmentation, state-based modelling, and unsupervised motif discovery, enabling classification of feeding, resting, grooming, and abnormal motor patterns. Integration with machine learning approaches for automated behaviour annotation offers a path toward richer, scalable phenotyping without increasing manual burden.

More broadly, platforms such as BEEhaviourLab align with a growing shift toward incorporating behavioural endpoints into ecotoxicological screening. Standardised, reproducible behavioural metrics across species and stressors create opportunities to compare sublethal effects on a common scale and to integrate behavioural disruption into next-generation risk assessment frameworks. By making subtle neurobehavioural impairment measurable at scale, automated phenotyping systems can help close the gap between laboratory toxicity testing and ecologically meaningful outcomes.

4. Methods

4.1. BEEhaviourLab equipment

BEEhaviourLab is a modular, distributed system for automated, high-throughput recording of insect behaviour under controlled experimental conditions. The system is designed to scale from a single recording unit to many parallel nodes, each operating as an independent BEEhaviourLab unit and coordinated through a central control computer (Figure 1 A-B).

Each node integrates synchronized video recording (Raspberry Pi NoIR v2 camera), audio capture (RØDE smartLav+), and environmental monitoring (temperature and humidity via a DHT22 sensor). Lighting is provided under the clear perspex recording arena using LED strip lights, with the option to use white LEDs (daylight) or red (night mode). The physical enclosure is constructed from laser-cut perspex and assembled using plastic fasteners or adhesive (Figure 1 A).

Each node is controlled by a Raspberry Pi 4B microcontroller (4 GB RAM). Experimental parameters are defined through a human-readable configuration file, allowing users to specify metadata (e.g. species, treatment, replicate), recording duration, inter-recording intervals, number of recording segments, and video and audio capture and compression settings (Figure 1 C). This design allows identical experimental protocols to be deployed across many nodes with minimal manual intervention.

Nodes are accessed remotely via SSH over Wi-Fi from a central node computer, which can initiate recordings, manage multiple BEEhaviourLab units in parallel, and coordinate data acquisition across the system. Recording is initiated by running a single control script on each node. Following each recording segment, video and audio data are compressed and saved together with structured metadata. Data are transferred from individual nodes to the central computer for storage and downstream processing.

All hardware designs and Raspberry Pi control software are fully open source and available on GitHub (BEEhaviourLab/BEEhaviourLab-apparatus). The software is intentionally modular, enabling users to adapt the system to their own experimental needs, including adding additional sensors, modifying acquisition parameters, or extending recording and control functionality.

4.2. Automated insect tracking

Insects in BEEhaviourLab video recordings were automatically detected and tracked using YOLO version 8 “nano” (YOLOv8n, Ultralytics), a lightweight object-detection architecture suitable for rapid fine-tuning and deployment on standard computing hardware. To support reuse and adaptation, we provide a step-by-step tutorial for retraining the model on user-defined insect datasets (GitHub: BEEhaviourLab/BEEhaviourLab-YOLO-training).

Training and validation data were derived from video frames extracted from BEEhaviourLab recordings. Frames were manually annotated in Label Studio (labelstud.io) with two object

classes: *bee* and *feeder*. Two YOLOv8n models were trained for 100 epochs: model 1 was trained on 155 annotated frames from recordings of bumble bees (*Bombus terrestris*); and model 2 was trained on a multi-species dataset comprising 455 annotated images across four insect species: honey bee (*Apis mellifera*), ivy bee (*Colletes hederae*), bumble bee (*Bombus terrestris*), and hoverfly (*Eristalis tenax*).

Initial tracking of multiple objects from frame to frame is performed using the BoT-SORT algorithm [34]. BoT-SORT is generally deployed in scenarios where there are an unknown number of objects to track and this number can change during the course of the video. The method incorporates measures to reassign identifications (IDs) to low detection scores, such as when a bee is occluded by another bee or by the feeder for a certain number of frames. Despite this, the tracking remained prone to errors when re-associating an ID with an individual and this often resulted in the assigned ID being incremented, as if a new individual had appeared, rather than an previous ID being reused. Since, in our case, we had a fixed number of individuals within the field of view, we added our own data processing step after BoT-SORT tracking. For each frame, detections are matched to a fixed set of stable identities (IDs 1…N) using the Hungarian algorithm [35], minimising Euclidean displacement from the previous frame's positions. Stable IDs therefore persist even when detections are temporarily missing: if an individual is not detected in a given frame, its last known position is propagated forward and an empty row is emitted. This enforces a constant population size and prevents ID creation or deletion.

Model performance was evaluated on validation datasets consisting of recordings from independent experimental cohorts, ensuring no overlap between training and evaluation data. Performance metrics included recall (proportion of true objects detected), precision (proportion of detections that are correct), mAP50 (mean average precision calculated using detections with at least 50% overlap between predicted and true bounding boxes), mAP50-95 (mean average precision averaged across overlap thresholds from 50% to 95%), and the F1-score (harmonic mean of precision and recall).

Bee locomotion was quantified using movement speed, which was calculated for each tracked individual and summarised for each recording by aggregating across bees within an experimental cage.

4.3. Audio analyses

Raw audio recordings were processed to detect active (flight or buzzing sound-containing) segments using a custom pipeline. Audio was resampled to 16 kHz and bandpass-filtered (200–8000 Hz, 4th-order Butterworth) to focus on bee-relevant frequencies. Frame energy (dB) was computed with 0.1 s windows and 0.05 s hops. A robust threshold (median + k×MAD, k = 5, or percentile-based for high-energy files) classified frames as containing buzzing sound or no sound. The resulting binary mask was cleaned by removing segments shorter than 0.1 s and filling gaps shorter than 0.15 s. For each 60 s audio file, we computed the proportion of recording time containing buzzing sounds.

4.4. Tracking performance across insect species

To evaluate the flexibility of the automated tracking pipeline across taxa, four insect species were recorded in BEEhaviourLab: honey bees (*Apis mellifera*, n=21), ivy bees (*Colletes hederae*, n=3), and hoverflies (*Eristalis tenax*, n=5). Insects were wild-caught immediately prior to recording, and were released following the recording period, and were housed with 1 bee per cage. We additionally recorded from lab-kept bumble bees (*Bombus terrestris,* n=33) purchased from BioBest (Westerlo, Belgium, via Agralan, Swindon, UK). Data was recorded with 12 hrs light (06:00 - 18:00) and 12 hrs dark (night, 18:00 to 06:00).

Recordings consisted of 5 min video segments collected at 25 min intervals over a 48 h period, yielding a time series of 96 recordings per cage. These data were used to assess detection and tracking performance across species with differing body sizes, morphologies, and movement patterns.

4.5. Moxidectin toxicity in bumble bees

To demonstrate the application of BEEhaviourLab for sublethal toxicity assessment, we quantified behavioural effects of moxidectin, a broad-spectrum antiparasitic compound used in veterinary and human medicine [36], on Bombus terrestris over a 48 h period following contact exposure.

Individual bees received a 1 µl droplet applied to the dorsal thorax containing moxidectin (Cydectin Pour On formulation, 0.5% w/v) diluted in DMSO. Exposure doses were 0 (vehicle control; n = 30), 0.782 (n = 33), 7.82 (n = 36), 78.2 (n = 33), 782 (n = 30), and 7,820 (n = 45) ng moxidectin per bee. Following exposure, bees were placed into clear perspex recording cages (three bees per cage) and provided with 1 M sucrose via a syringe feeder. Cages were loaded into BEEhaviourLab nodes and recorded for 1 min segments at 14 min intervals over 48 h, resulting in 192 recordings per cage. Data was recorded with 12 hrs light (06:00 - 18:00) and 12 hrs dark (night, 18:00 to 06:00). All bees exposed to the highest dose (7,820 ng per bee) were excluded from subsequent analyses due to complete mortality within 1–12 h.

At the conclusion of the recording period, experimental metadata were collected for each cage, including total sucrose solution consumption, individual body mass, and mortality assessed at the cage level. All surviving bees were euthanised by placing the cages in the freezer overnight.

4.6. Statistical analyses

Statistical analyses were performed in R Studio [37]. Behavioural data were analysed using generalized linear mixed-effects models implemented in glmmTMB [38], with cage identity included as a random intercept in all models to account for repeated measures. Model diagnostics were assessed using simulation-based residual checks in DHARMa [39] and model-based estimates and contrasts were obtained using emmeans [40].

Locomotion speed was analysed using a two-part mixed-effects framework separating behavioural state from movement intensity. Individual bees were classified as moving or inactive using a data-derived speed threshold based on the low-speed mode of the speed distribution.

The probability of movement was modelled using a binomial GLMM with a logit link, while locomotion speed conditional on movement was modelled using a Gamma GLMM with a log link. For the bumble bee dose–response experiment, log-transformed dose and time were included as covariates; for the wild bee dataset, species and time were included as fixed effects.

Circadian patterns of locomotor activity were analysed for both datasets using harmonic mixed-effects models, in which time was represented by 24-h sine and cosine terms to capture continuous circadian periodicity. Movement probability was modelled using binomial GLMMs including species or dose and interactions with the harmonic terms. The presence and modulation of circadian rhythmicity were assessed using likelihood-ratio tests of the harmonic interaction terms. Social spacing (mean inter-bee distance) and exploration were analysed using Gamma mixed-effects models with log-transformed dose and time as covariates.

Audio-derived behavioural measures were analysed using zero-inflated beta generalized linear mixed models, which explicitly separate the probability of complete inactivity from the distribution of non-zero proportions. Dose (log-transformed) and time were included as covariates, with cage identity included as a random intercept. Model selection was based on AIC, and inference was based on model-based marginal means and contrasts.

To synthesise multimodal effects into a single sublethal toxicity metric, we computed cycle-averaged predicted differences from control for each behavioural component and dose, averaged over the circadian cycle. These dose-specific effects were scaled by the between-individual standard deviation observed in controls, yielding an sublethal toxicity index expressed in units of control behavioural variability (standard deviation, SD). A behavioural effect threshold was defined at one control SD, such that index values ≥1 indicate behavioural changes comparable in magnitude to natural between-individual variability in untreated bees. This threshold was used to estimate the dose at which sublethal behavioural effects first appear.

5. Funding

Funding for this project was provided by Schmidt Sciences through the Schmidt AI in Science Fellowship awarded to R.H.P. and C.N., and by the Crankstart Foundation through the Crankstart Internship Programme awarded to A.S. Additional support for J.C.Z.K was provided by the Royal (Dick) School of Veterinary Studies.

6. Figures

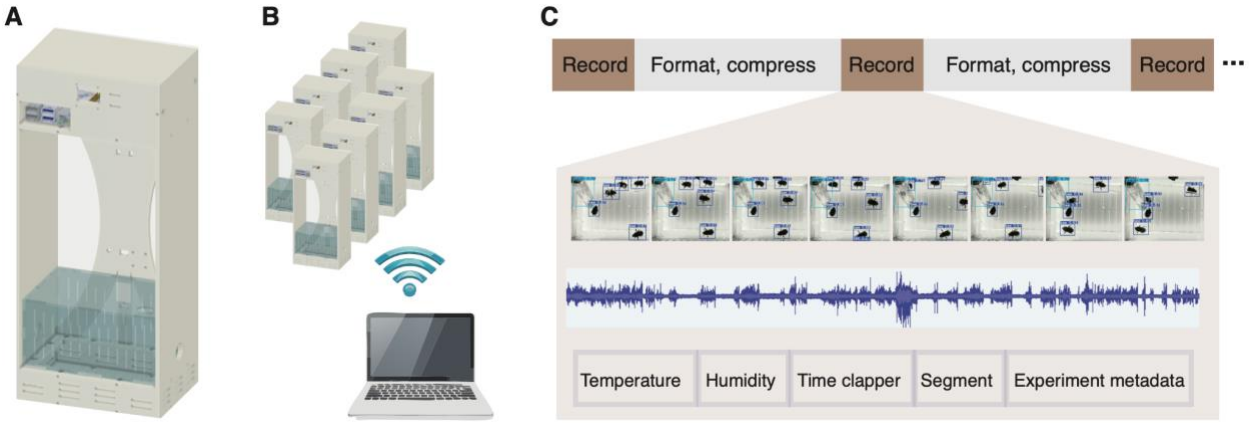

Figure 1: High-throughput insect behavioural tracking using BEEhaviourLab.

(A) Schematic of a single BEEhaviourLab unit, showing the Perspex enclosure (white) and transparent recording arena (clear). (B) Multiple units operate in parallel and communicate wirelessly with a central node computer. (C) User-defined recording protocols enable flexible specification of recording duration (brown), inter-recording intervals (grey), and number of recording segments. During inter-recording intervals, video, audio, sensor data, and experimental metadata are formatted, compressed, and transferred to the central node via Wi-Fi.

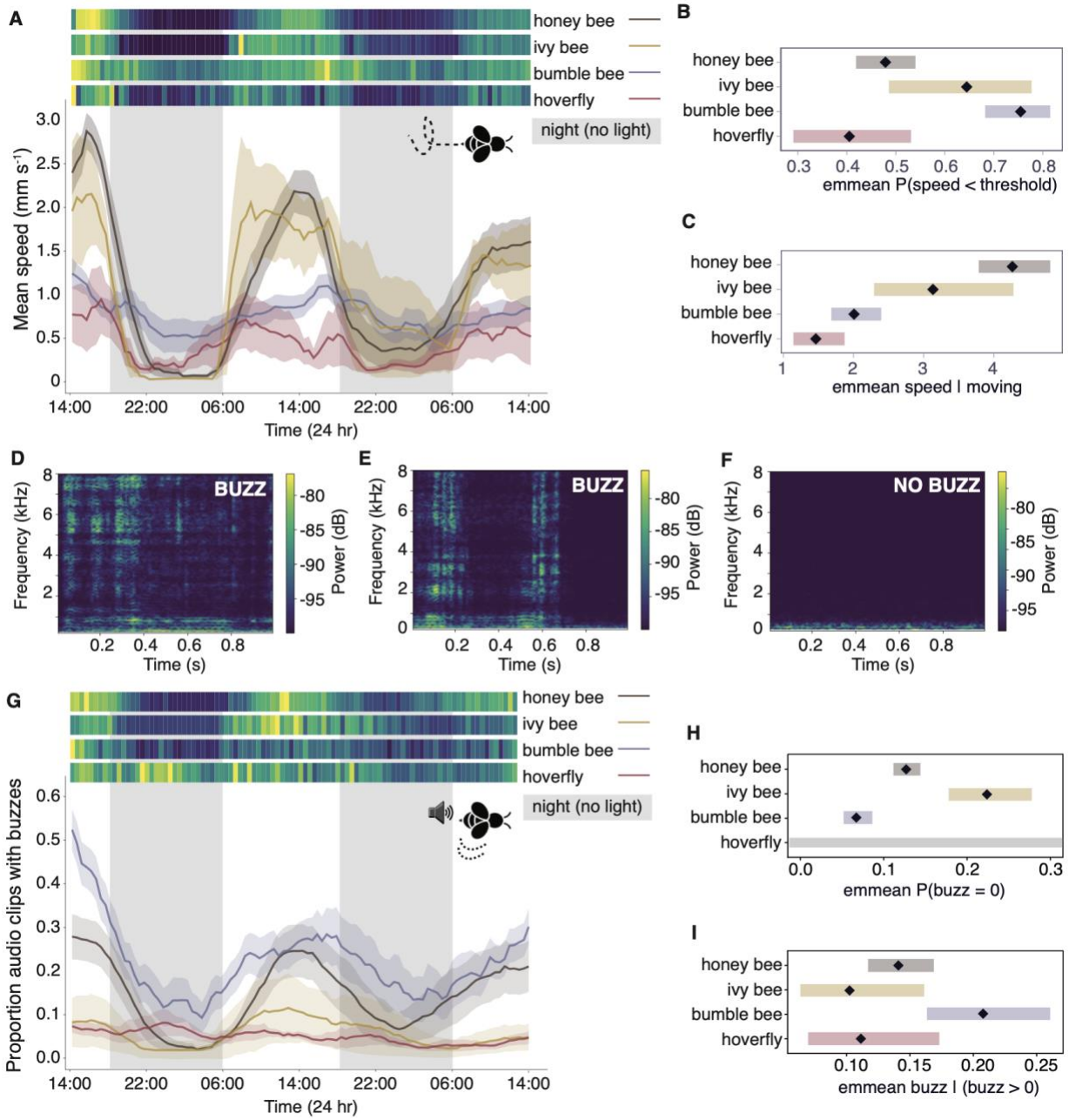

Figure 2: Cross-species locomotor and acoustic activity patterns.

(A) Locomotion speed over 48 h by species; heatmaps are scaled to species-specific maxima, with lines showing mean ± SD across individuals. (B) Estimated marginal means (EMMs) for probability of movement. (C) EMMs for locomotion speed conditional on movement. (D–F) Example spectrograms showing audio clips with buzzing (D–E) or no buzz (F). (G) Proportion of clips containing buzzes over 48 h by species; heatmaps scaled to species-specific maxima, lines show mean ± SD. (H) EMMs for probability of no buzzing (hoverflies excluded due to absence of complete zeros). (I) EMMs for proportion of clips with buzzes, conditional on non-zero observations.

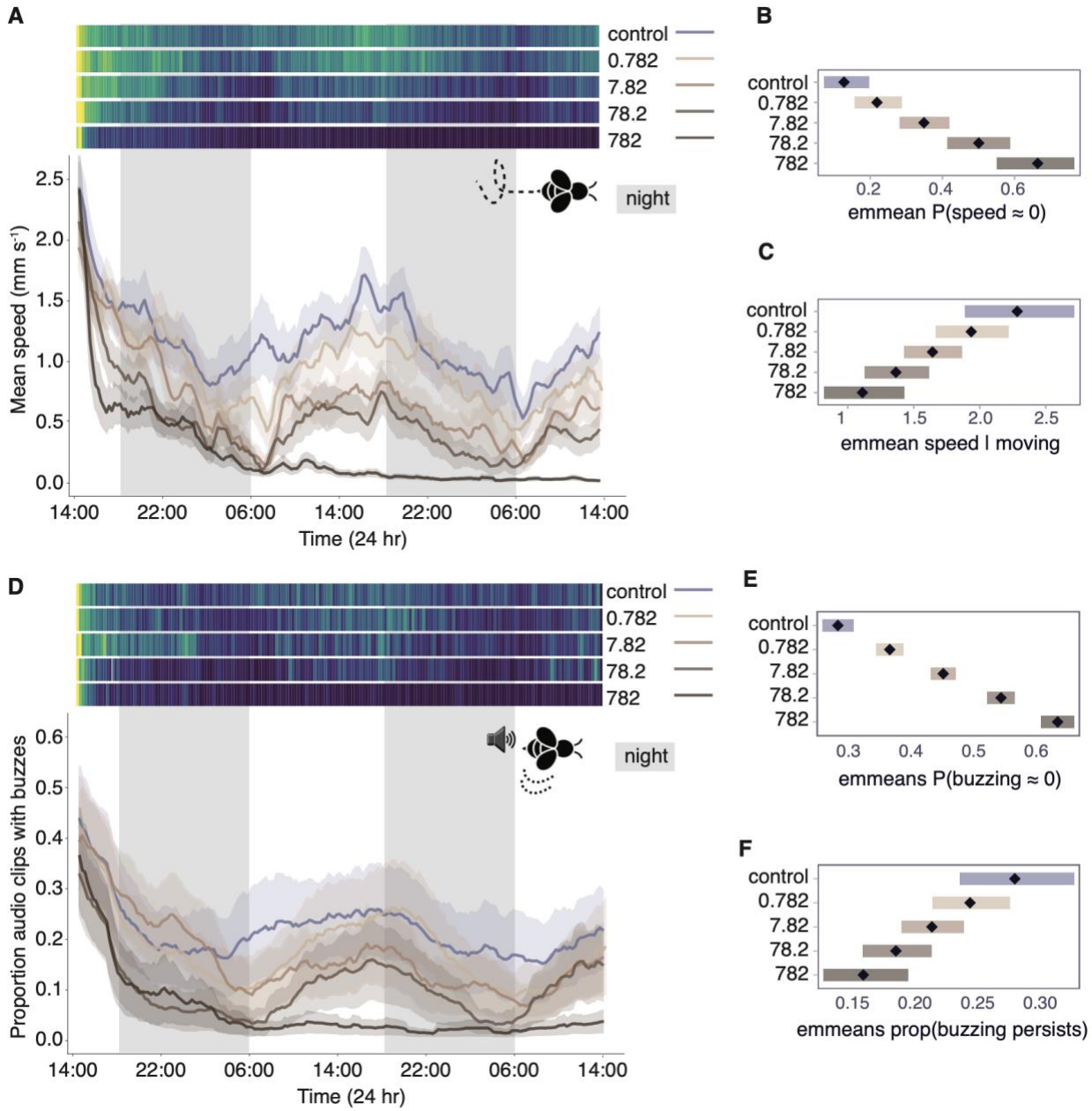

Figure 3: Moxidectin induces dose-dependent reductions in locomotion and buzzing.

(A) Mean locomotion speed over 48 h across moxidectin treatments; heatmaps are scaled to the maximum within each treatment group. (B) Estimated marginal means (EMMs) for the probability that speed falls below the movement threshold, showing dose-dependent increases in inactivity. (C) EMMs for locomotion speed conditional on movement, indicating more modest dose-dependent reductions among active bees. (D) Mean proportion of 1 s clips containing buzzing over 48 h across treatments; heatmaps scaled to treatment-specific maxima. (E) EMMs for the probability of no buzzing, showing dose-dependent increases in complete inactivity. (F) EMMs for buzzing persistence conditional on buzzing occurring, indicating smaller reductions once buzzing is initiated.

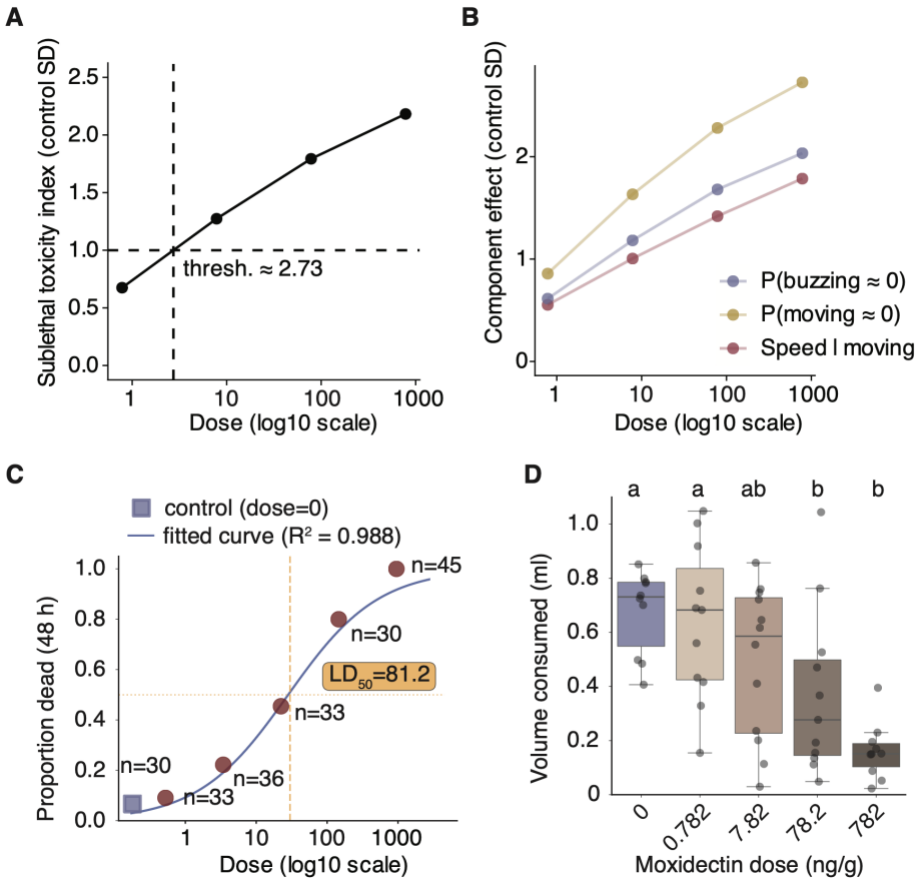

Figure 4. A multimodal behavioural toxicity index detects sublethal effects below lethal thresholds.

A: Sublethal toxicity index integrating deviations in locomotion and buzzing relative to control variability. The dashed line indicates a threshold of functional abnormality at one control standard deviation, crossed at 2.73 ng bee^{-1} moxidectin. B: Relative contributions of individual behavioural components to the composite toxicity index, showing dominance of changes in behavioural engagement over graded performance effects. C: 48 h contact mortality curve with an estimated LD_{50} of 81.2 ng bee^{-1} moxidectin. D: Total sugar syrup consumption per cage (three bees per cage) across control and moxidectin treatments. Letters denote significant differences between doses based on estimated marginal means.

Table 1: YOLO model benchmarking across bee species.

Model	Species	N validation images	Precision	Recall	F1-Score	mAP50	mAP50-95
model 1	bumble bee	50	0.98	0.97	0.98	0.98	0.77
	ivy bee	38	0.50	0.45	0.47	0.47	0.37
	honey bee	40	0.76	0.67	0.74	0.74	0.39
	hoverfly	38	0.83	0.67	0.75	0.75	0.36
model 2	bumble bee	50	0.98	0.98	0.98	0.98	0.85
	ivy bee	38	0.99	0.99	0.99	0.99	0.85
	honey bee	40	1.00	1.00	1.00	1.00	0.84
	hoverfly	38	1.00	0.99	0.99	0.99	0.64

References

1. Chen D, Huang H, Huang Y, Yang W, Shan W, Hao G, et al. Toxicity Tests for Chemical Pesticide Registration: Requirement Differences among the United States, the European Union, Japan, and China? J Agric Food Chem. 2023;71: 7192–7200. doi:10.1021/acs.jafc.3c00410

2. Rohr JR, Salice CJ, Nisbet RM. The pros and cons of ecological risk assessment based on data from different levels of biological organization. Crit Rev Toxicol. 2016;46: 756–784. doi:10.1080/10408444.2016.1190685

3. Müller C. Impacts of sublethal insecticide exposure on insects — Facts and knowledge gaps. Basic Appl Ecol. 2018;30: 1–10. doi:10.1016/j.baae.2018.05.001

4. Maggi T, Pardo L, Chreil R. Pollinator Diversity: A Key to Ecosystem Resilience and Food Security. 2023.

5. Van Der Sluijs JP, Vaage NS. Pollinators and Global Food Security: the Need for Holistic Global Stewardship. Food Ethics. 2016;1: 75–91. doi:10.1007/s41055-016-0003-z

6. Botías C, David A, Hill EM, Goulson D. Quantifying exposure of wild bumblebees to mixtures of agrochemicals in agricultural and urban landscapes. Environ Pollut. 2017;222: 73–82. doi:10.1016/j.envpol.2017.01.001

7. Mahefarisoa KL, Simon Delso N, Zaninotto V, Colin ME, Bonmatin JM. The threat of veterinary medicinal products and biocides on pollinators: A One Health perspective. One Health. 2021;12: 100237. doi:10.1016/j.onehlt.2021.100237

8. Thompson DA, Lehmler H-J, Kolpin DW, Hladik ML, Vargo JD, Schilling KE, et al. A critical review on the potential impacts of neonicotinoid insecticide use: current knowledge of environmental fate, toxicity, and implications for human health. Environ Sci Process Impacts. 2020;22: 1315–1346. doi:10.1039/C9EM00586B

9. Zioga E, Kelly R, White B, Stout JC. Plant protection product residues in plant pollen and nectar: A review of current knowledge. Environ Res. 2020;189: 109873. doi:10.1016/j.envres.2020.109873

10. Wolstenholme AJ, Rogers AT. Glutamate-gated chloride channels and the mode of action of the avermectin/milbemycin anthelmintics. Parasitology. 2005;131: S85–S95. doi:10.1017/S0031182005008218

11. Wolstenholme AJ. Glutamate-gated Chloride Channels. J Biol Chem. 2012;287: 40232–40238. doi:10.1074/jbc.R112.406280

12. Suárez VH, Lifschitz AL, Sallovitz JM, Lanusse CE. Effects of faecal residues of moxidectin and doramectin on the activity of arthropods in cattle dung. Ecotoxicol Environ Saf. 2009;72: 1551–1558. doi:10.1016/j.ecoenv.2007.11.009

13. Verdú JR, Cortez V, Martinez-Pinna J, Ortiz AJ, Lumaret J-P, Lobo JM, et al. First assessment of the comparative toxicity of ivermectin and moxidectin in adult dung beetles: Sub-lethal symptoms and pre-lethal consequences. Sci Rep. 2018;8: 14885. doi:10.1038/s41598-018-33241-0

14. Abdu-Allah GAM, Pittendrigh BR. Lethal and sub-lethal effects of select macrocyclic lactones insecticides on forager worker honey bees under laboratory experimental conditions. Ecotoxicology. 2018;27: 81–88. doi:10.1007/s10646-017-1872-6

15. Peterson EM, Green FB, Smith PN. Toxic responses of blue orchard mason bees (Osmia lignaria) following contact exposure to neonicotinoids, macrocyclic lactones, and pyrethroids. Ecotoxicol Environ Saf. 2021;208: 111681. doi:10.1016/j.ecoenv.2020.111681

16. Peterson EM, Thompson KN, Shaw KR, Tomlinson C, Longing SD, Smith PN. Use of nest bundles to monitor agrochemical exposure and effects among cavity nesting pollinators. Environ Pollut. 2021;286: 117142. doi:10.1016/j.envpol.2021.117142

17. Dirilgen T, Herbertsson L, O'Reilly AD, Mahon N, Stanley DA. Moving past neonicotinoids and honeybees: A systematic review of existing research on other insecticides and bees. Environ Res. 2023;235: 116612. doi:10.1016/j.envres.2023.116612

18. Baatrup E, Bayley M. Animal Locomotor Behaviour as a Health Biomarker of Chemical Stress. In: Seiler JP, Autrup JL, Autrup H, editors. Diversification in Toxicology — Man and

Environment. Berlin, Heidelberg: Springer Berlin Heidelberg; 1998. pp. 164–178. doi:10.1007/978-3-642-46856-8_15

19. Bownik A. Daphnia swimming behaviour as a biomarker in toxicity assessment: A review. Sci Total Environ. 2017;601–602: 194–205. doi:10.1016/j.scitotenv.2017.05.199

20. Cassar S, Adatto I, Freeman JL, Gamse JT, Iturria I, Lawrence C, et al. Use of Zebrafish in Drug Discovery Toxicology. Chem Res Toxicol. 2020;33: 95–118. doi:10.1021/acs.chemrestox.9b00335

21. Williamson SM, Willis SJ, Wright GA. Exposure to neonicotinoids influences the motor function of adult worker honeybees. Ecotoxicology. 2014;23: 1409–1418. doi:10.1007/s10646-014-1283-x

22. Dong B, Chen W. A high precision method of segmenting complex postures in Caenorhabditis elegans and deep phenotyping to analyze lifespan. Sci Rep. 2025;15: 8870. doi:10.1038/s41598-025-93533-0

23. O'Brien TJ, Barlow IL, Feriani L, Brown AE. High-throughput tracking enables systematic phenotyping and drug repurposing in C. elegans disease models. eLife. 2025;12: RP92491. doi:10.7554/eLife.92491

24. Zhao X, Li Z, Cao B, Jin Y, Wang W, Tian J, et al. A high-throughput system for drug screening based on the movement analysis of zebrafish. Heliyon. 2024;10: e36495. doi:10.1016/j.heliyon.2024.e36495

25. Jordi J, Guggiana-Nilo D, Bolton AD, Prabha S, Ballotti K, Herrera K, et al. High-throughput screening for selective appetite modulators: A multibehavioral and translational drug discovery strategy. Sci Adv. 2018;4: eaav1966. doi:10.1126/sciadv.aav1966

26. Bedrossiantz J, Martínez-Jerónimo F, Bellot M, Raldua D, Gómez-Canela C, Barata C. A high-throughput assay for screening environmental pollutants and drugs impairing predator avoidance in Daphnia magna. Sci Total Environ. 2020;740: 140045. doi:10.1016/j.scitotenv.2020.140045

27. Jain A, Tang ATS, Crall J. Ethomics for Ecotoxicology: Automated Tracking Reveals Diverse Effects of Insecticides on Bumble Bee Foraging and In-Nest Behavior. Integr Comp Biol. 2025;65: 919–931. doi:10.1093/icb/icaf118

28. Easton-Calabria A, Crall JD. The BumbleBox: An open-source platform for quantifying behaviour in bumblebee colonies. Ecol Solut Evid. 2025;6: e70052. doi:10.1002/2688-8319.70052

29. Geissmann Q, Garcia Rodriguez L, Beckwith EJ, French AS, Jamasb AR, Gilestro GF. Ethoscopes: An open platform for high-throughput ethomics. PLOS Biol. 2017;15: e2003026. doi:10.1371/journal.pbio.2003026

30. Jones H, Willis JA, Firth LC, Giachello CN, Gilestro GF. A reductionist paradigm for high-throughput behavioural fingerprinting in Drosophila melanogaster. eLife. 2023;12: RP86695. doi:10.7554/eLife.86695

31. Laurent F, Blanc A, May L, Gándara L, Cocanougher BT, Jones BMW, et al. LarvaTagger: manual and automatic tagging of *Drosophila* larval behaviour. Peng H, editor. Bioinformatics. 2024;40: btae441. doi:10.1093/bioinformatics/btae441

32. Gandara L, Jacoby R, Laurent F, Spatuzzi M, Vlachopoulos N, Borst NO, et al. Pervasive sublethal effects of agrochemicals on insects at environmentally relevant concentrations. Science. 2024;386: 446–453. doi:10.1126/science.ado0251

33. Tasman K, Hidalgo S, Zhu B, Rands SA, Hodge JJL. Neonicotinoids disrupt memory, circadian behaviour and sleep. Sci Rep. 2021;11: 2061. doi:10.1038/s41598-021-81548-2

34. Aharon N, Orfaig R, Bobrovsky B-Z. BoT-SORT: Robust Associations Multi-Pedestrian Tracking. arXiv; 2022. doi:10.48550/arXiv.2206.14651

35. Kuhn HW. The Hungarian method for the assignment problem. Nav Res Logist Q. 1955;2: 83–97. doi:10.1002/nav.3800020109

36. Schraven AL, Stannard HJ, Old JM. A systematic review of moxidectin as a treatment for parasitic infections in mammalian species. Parasitol Res. 2021;120: 1167–1181. doi:10.1007/s00436-021-07092-0

37. Posit team. RStudio: Integrated Development Environment for R. Boston, MA: Posit Software, PBC; 2025. Available: http://www.posit.co/}

38. Williams C, McGillycuddy M, Drobniak SM, Bolker BM, Warton DI, Nakagawa S. Fast phylogenetic generalised linear mixed-effects modelling using the glmmTMB R package.

39. Hartig F. DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models. 2024. doi:10.32614/CRAN.package.DHARMa

40. Lenth RV, Piaskowski J. emmeans: Estimated Marginal Means, aka Least-Squares Means. 2026.