Biology-Informed inverse problems for insect pests detection using pheromone sensors

Thibault Malou[1,2], Nicolas Parisey[2], Katarzyna Adamczyk-Chauvat[1], Elisabeta Vergu[1], Beatrice Laroche[1], Paul-Andre Calatayud[3], Philippe Lucas[4], and Simon Labarthe[5,6]

[1]Université Paris-Saclay, INRAE, MaIAGE, 78350, Jouy-en-Josas, France
[2]INRA, Institute of Genetics, Environment and Plant Protection (IGEPP—Joint Research Unit 1349), Le Rheu, France
[3]Institut de Diversité, Ecologie et Evolution du Vivant (IDEEV), Université Paris-Saclay, CNRS, IRD, UMR Evolution, Génome, Comportement et Ecologie (EGCE), Gif-sur-Yvette, France
[4]Institute of Ecology and Environmental Sciences, INRAE, Versailles, France
[5]Univ. Bordeaux, INRAE, BIOGECO, 33610, Cestas, France
[6]Inria, INRAE, Pléiade, 33400, Talence, France

Dear Recommender, dear Reviewers,

First of all, we would like to warmly thank you for your constructive reviews of our first submission. We believe that they have helped us to considerably improve the quality of our paper.

We have carefully taken into account all your comments. You will find below the different answers to your questions and the way we included your remarks in the manuscript.

We would also like to disclose that, during the revision process, we found a bug in the code of the computation of the adjoint model that had a slight impact on the resolution of the BI-DA model. We re-computed the whole set of computations, but their interpretation, and therefore the conclusions of the paper,remained unchanged.

We hope that this new version of the paper will meet your expectations.
Best regards
Simon Labarthe, for the authors.

1 Recommender

The modifications corresponding to Reviewer 1 are indicated in magenta in the revised manuscript.

1. **Reviewer comment:** The submitted paper proposes a data assimilation framework for an inverse problem. The inverse problem consists in predicting an insect presence landscape given a network of pheromone sensors and a pheromone diffusion model. The data assimilation framework consists in using the control term of the inverse problem to guide the optimisation by some biological knowledge of specific pests, such as some parameters of population dynamics (BI-DA). The motivation is a more rational use of pesticides than prophylaxis or less informed models. The accuracy and precision of the results are tested on a simple toy example to test the ability of the model and the algorithm to detect the source of the pheromones and the efficiency of the BI-DA principle. A further simulation is then carried out on a real plot with realistic parameters and regulations based on knowledge of a maize pest. A repositioning of the sensors (informed by the data of the initial positions) is carried out during the test phase to allow a better detection. The paper has been well evaluated by two referees who, based on their comments, propose to ask for a revision. Their comments should be useful to improve the paper and should not imply major changes. I also have a list of suggestions, which are not mandatory and should be considered as a guide to improve the paper if the authors find them relevant.
   Answer: We would like to thank you really much for this positive comment.

2. **Reviewer comment:** The introduction could be improved by some minor clarifications of the motivation. From line 44, it is argued that the use of pesticides when pests are detected could be an economy compared to prophylaxis, and implicitly that knowing the location of insects can lead to additional economies. Is it because pesticides are used less frequently (only when pests are detected), or because they are used in smaller quantities (based on an estimate of a population size), or because they can only be used in parts of fields where insects are detected? Are there empirical studies showing that the use of sensors actually leads to less pesticide use than prophylaxis, and that the use of more precise insect positions through inverse problems leads to less pesticide use than the presence/absence signal from sensors?
   Answer: In a precision farming context, earlier and more accurate prediction of insect presence could result in pesticide use reduction thanks to the three levers you mentioned. Thanks to detection-guided pesticide use, treatment frequency, dose and surface could be reduced. To our knowledge, there is still no empirical studies indicating that using sensors leads to less pesticide use than prophylaxis.

We reformulated the introduction to make more explicit these three ways of reducing pesticide use.
Modifications: We have made a modification line 49.

3. **Reviewer comment:** The mathematical model deals with the concentration of a particular pheromone, specific to a particular insect. The introduction says that the sensors can detect many different types of pheromone. So is there one model per insect species? When used in the field, are there many models running in parallel? What is the typical number of models?
Answer: We clarified this point in the discussion, as part of a new section titled *Toward an effective use of BI-DA for pheromone sensor analysis.*
Modifications: See modifications line 949.

4. **Reviewer comment:** The running time of the optimisation could be compared with the insects' breeding time to know if the real time application is feasible.
Answer: We addressed this issue in the new section of the discussion *Toward an effective use of BI-DA for pheromone sensor analysis.* In the present implementation of BI-DA, computation times may be in the order of magnitude of infestation settlement (i.e. tens of hours), which would impair BI-DA use in real context. However, surrogate models can be derived to speed-up computations in order to overcome this limitation.
Modifications: see line 936

5. **Reviewer comment:** - I did not understand part 3.4.4 and the statement that some sensors could be removed. I understood that this means knowing the location of the insects, or at least the signal from the sensors, but how can they be removed if this is unknown a priori? Besides this point, a sensitivity test on the number of sensors would be informative, did you choose the number of sensors in the test with an intuition of an optimal number?
Answer: In part 3.4.4, we quantify the total contribution of each sensor to the reconstruction of s_a, the optimal emitter distribution. To this end, we introduce the metric $\|c^*_{i,tot}\|_\infty$ that recapitulates the impact of a given sensor i on the decrease of the cost function. This metric allows to rank sensors by contribution and to detect sensors with marginal impact on optimization. Therefore, computing this metric gives a proxy of the number of sensors needed for accurate estimation in a next run (i.e. total number of sensors minus the number of sensors with negligible contribution). Optimal placement design is another strategy to select the number of sensors to be deployed in a given landscape.
Modifications: In line 682 part 3.4.4, we rephrased the description of use of $\|c^*_{i,tot}\|_\infty$ to make it more explicit. The link between optimal placement design and sensor number selection has been highlighted in line 1039,

6. **Reviewer comment:** - I sometimes got lost in the description of the results for the toy example test. It is long, has a lot of information and

I hardly got the most important ones (like "in most cases the "all-reg" performs better than the "no reg""). From this part I got the impression that BI-DA has little effect and that this depends strongly on the combination of the regulations, together with some parameters of the simulation. If I am wrong, there should be a sentence summarising the results. As the article is quite long, you might consider shortening this part.
Answer: We substantially shortened the result part for the toy case (around 30 %) and simplified the descriptions. We also provided for each section a short summary of the main findings, in order to ease the reading.
Modifications: Since modifications were too numerous, we did not highlighted text reductions.

Conclusive sentences have been added lines 605, 629, 675 and 709.

7. **Reviewer comment:** The use of sensor repositioning would be mentioned in the introduction or in the methods. It seems important for the understanding of the results and seems both a drawback of the method (if real repositioning is needed in the field) and a methodological contribution (although you admit in the conclusion that it is beyond this work).
Answer: Thank you for highlighting that this point should be better introduced. We estimated that this point is indeed beyond this work. Moreover, we are currently working on sensor positioning and repositioning. Thus, we wanted to be succinct on this matter in this study. You are right though that it should be better introduced to ease the understanding of the FAW test case and its results.
Modifications: see line 605

8. **Reviewer comment:** The description of the results of the FAW test case is also very long, describing some tiny details of the figure and missing an interpretation message.
Answer: We also strongly reduced the description of the different results, and added a take-home message for each paragraph.
Modifications: We added interpretation messages lines 821, 840 and 851.

9. **Reviewer comment:** As the classical data assimilation (DA) problem seems to be ill-conditioned and allows to reconstruct the the amount of pheromone emitted in the vicinity of the sensors, we propose". I did not understand this sentence: does "ill-conditioned" have a definition? Why "seems" and not "is"? I did not see the articulation.
Answer: By ill-conditionned data assimilation problem we meant that the Hessian matrix of the cost function is ill-conditionned. Eventhough we do not use optimization methods based on the Hessian and its inverse (such as Newton method), having a ill-conditionned Hessian matrix is an issue as it means that the cost function is “flat” (big variations in some component of the control variable leads to very small variations in the cost function) which make the optimization more complicated and requires

low learning rate (common problem in deep neural network learned with stochastic gradient descent).

Up to now, in the present study, we do not have any proof that the Hessian is ill-conditioned as the eigenvalues were not computed and cannot be computed in a decent computation time due to the size of the test cases. However, our expertise and the different numerical experiments we performed highlight that the cost function is rather flat (in the sense given earlier).

We decided to remove the word "ill-conditioned" and to reformulate this sentence.
Modifications: see line 864

10. **Reviewer comment:** In the discussion you mention a possible multiple repositioning of sensors, but I don't understand how this can be related to reality in a field. Is it possible to change the position of sensors several times during a pest invasion?
    Answer: As shown in this study, the physical scales involved in pheromone propagation are such that a dense spatial sampling is needed for accurate pest detection. This dense sampling can be obtained by a large network of fixed sensors, or by a lower number of mobile sensors. We chose to dig into the hypothesis of mobile devices because 1) it allowed to work with the hypothesis of a limited budget of sensors, 2) it gave us the opportunity to develop the one-sensor criteria enabling an in-line detection of less informative sensors, 3) mobility can be added to the technical specifications of the sensors that are still in development.
    Modifications: A sentence has been added in the discussion line 1020.

11. **Reviewer comment:** I wonder how much the conclusions depend on the specific design of the test cases. Some parameters have been arbitrarily fixed, do you have any intuition about the robustness of the conclusions to other types of designs?
    Answer: We agree that the sources of uncertainties are multiple in the inference process. We discussed them in Sec. 4.4. We extended it in order to discuss a systematic sensitivity analysis of the different parameters involved in the process. We felt that a systematic uncertainties quantification was too ambitious for this study, but will be the object of future works. We already explored during a master thesis project the impact of meteorological uncertainties on pheromone propagation and sensor signals.
    Modifications: Additional discussions have been added line 1001.

12. **Reviewer comment:** I also wondered about the distance of such a study from the application that motivates it. What would be the next steps towards realising the economic proposition?
    Answer: This question has been treated in the new section titled *Toward an effective use of BI-DA for pheromone sensor analysis.*
    Modifications: see lines 925 and 1024

13. **Reviewer comment:** I also wondered about the amount of computation needed to make the predictions, the pressure on resources to do it, and whether the balance of benefits and risks was so clearly in favour of using it to reduce pesticides.
    Answer: Computation times and possible methods that can be used to reduce them have been addressed in response to comment 1.3 above (see adding line 936). Here, a different question is raised: the balance of benefits and risks between BI-DA on the one hand and classical prophylaxis using pesticides on the other. Using BI-DA induces different environmental drawbacks, such as sensor fabrication, sensor deployment and computations. A global life cycle analysis (LCA) of the sensors should be conducted in order to measure the pressure on resource and the carbon release induced by their deployment, and compared to pesticide LCA. Comparison should also include the impact of both practices on health, water and biodiversity.
    Modifications: A discussion has been added line 930.

14. **Reviewer comment:** Possible typos:
    - insect pheromones are specie-specific : species-specific?
    - the same homogeneous diffusion tensor than the toy case is used : than in the toy case? Or as in the toy case?
    - In a first time we optimize : As a first step?
    - improves a little : slightly improve?
    - how a specific sensor impact : impacts
    - no insect were correctly located : no insect was?
    - that are not consider : considered?
    - following the following steps : using the following steps?

    Answer: Thank you for highlighting these typos.
    Modifications: see respectively lines 13, 273, 313, 525, 685, 756, 1006 and 1007

2 Reviewer 1

The modifications corresponding to Reviewer 2 are indicated in green in the revised manuscript.

1. **Reviewer comment:** The paper entitled: Biology-informed inverse problems for insect pests detection using pheromone sensors deals with a source estimation problem for a reaction-difusion-advection model in the context of insect invasion in agriculture. The setup is a standard Variational formulation for the source detection inverse problem. The main contribution consists in adding regularisation terms which enforce weakly (by penalisation) some expected behaviours. The paper is well written and presents a valuable contribution. It deserves publication after a revision. Please, find hereafter some comments and questions.

Answer: We thank Reviewer 1 for his/her positive comments.

2. **Reviewer comment:** 1) The 2d model of reaction-diffusion type is well justified in appendix B. In this model, the wind velocity is assumed to be constant in the x,y direction and negligible in the z direction. This motivates the fact that, in the z direction, the passive scalar falls due to gravity. It is assumed that u,v are near to average over the layer of interest. This could be a first reasonable assumption (depending on the size, in z, of the layer of interest). However, it is neglecting the ground boundary layer, which makes the components u,v to be depending on z. Could the authors:
   a) comment on the size in z of the layer of interest
   b) comment on this modelling assumption and on the impact it could have on the results?
   Answer: You are right that we approximate the vertical profile of the wind, including the ground boundary layer, with a simplified 2D model. This simplification is made due to the lack of available data on the vertical profile of the wind and to reduce the computational load.

   The size of the layer of interest is a few tens of centimetres thick, maximum one meter, and is located above the vegetation canopy. As our domains are located in open field with tall grass-like vegetation, the ground boundary layer should be rather thin with a small roughness length (few centimeters, that is below the vegetation canopy). We note that the effect of the ground boundary layer is not completely neglected in our model as the loss term specific to our 2D can be used to parametrize both the vertical fluxes and this vertical variations of wind velocity.

   As mentioned in the discussion about uncertainties and especially about the uncertainties related to the pheromone propagation model, more effort could be devoted in improving the model to take such uncertainties into consideration. One could for example think about training offline a 2D metamodel using a 3D CTM model as benchmark in order to tackle both the challenges of the computation time and of the uncertainties related to vertical variability averaged in a 2D model.

   This point and comment are improved in discussion section of article.
   Modifications: see lines 941, 988 and 994

3. **Reviewer comment:** 2) In the model, it is assumed that the initial concentration is zero. Is this modelling assumption justified? Could the authors comment on this point?
   Answer: This assumption is justified by the context of epidemiosurveillance: we put ourselves in the situation of pest-free fields that we want to protect from a primary invasion. Another justification is the emission pattern of the targeted pest, i.e. th FAW. First pheromone emissions take place during summer nights under a certain temperature. Thus, we set up the beginning of the time window before this first pheromone emission,

justifying the assumption that the initial concentration is zero.
Modifications: see line 123

4. **Reviewer comment:** 4) One of the main contributions of the paper consists in adding biology-informed regularisation terms. These consist in penalising the discrepancy with respect to some expected behaviour. Concerning the term in 3.1.3: a) it is a PDE residual based on a population dynamics model. What are the conditions for the residual to be in L^2?
   Answer: You are right that the population dynamics-informed regularization term is based on the L^2 norm of the residual of a population dynamic model and thus, we have to ensure that the residual is in L^2. We recall that the $j_{reg,PD}(s)$ term reads

   $$j_{reg,PD}(s) = \alpha_{reg,PD} \|\partial_t s - \frac{\partial_t q}{q} s + q\nabla \left(\sum_i \mathcal{F}_i(s) \right) - q \sum_i \mathcal{R}_i(s)\|^2_{L^2}.$$

   We first need that $log(q) \in C^1(0,T)$ so that $t \mapsto \frac{\partial_t q}{q}$ and $t \mapsto q$ are bounded on $(0,T)$. To check that $\left(\sum_i \mathcal{F}_i(s)\right) \in L^2(0,T;H^1(\Omega))$, that $\sum_i \mathcal{R}_i(s) \in L^2(0,T;\Omega)$ and that $\partial_t s \in L^2(0,T;\Omega)$, we need to ensure that $s \in \mathcal{S}$ where $\mathcal{S}$ is the functional space

   $$\mathcal{S} = \left(\bigcap_i \mathcal{F}_i^{-1} \left(L^2(0,T;H^1(\Omega)) \right) \right) \bigcap \left(\bigcap_i \mathcal{R}_i^{-1} \left(L^2(0,T;\Omega) \right) \right) \bigcap H^1(0,T;L^2(\Omega)).$$

   This functional space obviously depends on $\mathcal{F}_i$ and $\mathcal{R}_i$. In the Toy and FAW cases, the $\mathcal{F}_i$ are the null operators, and the mapping $\mathcal{R}_i$ is a smooth function so that $\mathcal{S} = H^1(0,T;L^2(\Omega))$.

   Next, we have to ensure that the sequence $(s_k)_{k\in\mathbb{N}}$ of the different steps of the optimization algorithm is in the space $\mathcal{S}$. Since

   $$s_{k+1} = s_k - \eta_k \nabla_s j(s_k) = s_k - \eta_k(\nabla_s j_{reg}(s_k) - c^*(s_k)),$$

   the regularity of the sequence depends on the regularity of the gradient of the cost function, that depends itself on the regularity of the adjoint state of the CTM and on the regularity of the adjoint of the population dynamic model operator. Here, the adjoint model of the CTM is a linear convection-diffusion PDE model with regular terms, so that the solutions of the adjoint model lies in $\mathcal{S}$. Furthermore, the adjoint of the population dynamic model operator is a set of regular ODE functions the solution of which are also in $\mathcal{S}$.
   Modifications: A remark has been added line 488.

5. **Reviewer comment:** b) Is the residual penalised in strong form? What about the boundary conditions for this model?
   Answer:

The residual is penalized in strong form. The boundary conditions are not explicitly taken into account in the penalties. If additional knowledge on boundary fluxes are obtained, a corresponding penalty can be added in the objective function of the minimization problem. Here, we do not make any assumptions on the boundaries and do not take them into consideration.
Modifications: A corresponding comment has been added line 488

6. **Reviewer comment:** 4) The authors discuss about the fact that the noise could induce negative source values, which are unphysical. Could a constraint of positive source be enforced?
   Answer: We agree with you: solving the minimization problem under positivity constraints could bypass this problem.
   Modifications: We added a comment line 563.

7. **Reviewer comment:** 5) In the discussion, several sources of uncertainty are described. Could the authors prospect on the possibility of performing Uncertainty Quantification and estimate the impact of the uncertainties on the results?
   Answer: This comment can be related to the comment 1.10 of the Recommender.
   Modifications: We added a comment line 1001.

3 Review by Angelo Iollo

The modifications corresponding to Reviewer 3 are indicated in orange in the revised manuscript.

1. **Reviewer comment:** Review by Angelo Iollo, 15 Oct 2024 17:04 The work is rich in content, rigorous, well-argued, and the numerical examples are well-documented. The main contribution lies in the formulation and comparison of different regularization terms for a data assimilation problem. The article undoubtedly deserves to be published. However, the authors could consider the following points, which may help clarify certain aspects:

   Answer: We thank Angelo Iollo for his positive evaluation.

2. **Reviewer comment:** 1. If I understand correctly, the control problem is posed in infinite dimensions, meaning that the number of degrees of freedom increase as the grid is refined. However, it is not clear which specific control variables are actually used and what the dimension of the control space is in the applied cases.
   Answer: Indeed the control problem is posed in infinite dimensions and when the spatio-temporal domain is discretized, the resulting discretized control variable has as many degree of freedom as the spatio-temporal grid. This point has been clarified in the manuscript, together with explicit

numbers of degrees of freedom.
Modifications: see lines 139, 142, 270 and 318

3. **Reviewer comment:** 2. In relation to point 1. it might be useful to precondition the search space to avoid infinite dimensionality, which could introduce high-frequency spatial instabilities and reduce the convergence rate.
   Answer: Thank you for highlighting that infinite dimensionality might raise such instabilities. We did not observe such instabilities in our situation, but it might be hidden by some other low convergence rate issue. In any case, preconditionning, such as random preconditionning or covariance-based preconditionning, might indeed be useful, for real applications, to improve the convergence rate and consequently the computation time. This point is added to the discussion on how to get to real applications, especially in terms of computation time reduction.
   Modifications: see line 943

4. **Reviewer comment:** 3. The authors emphasize that they use a gradient of the discrete problem. It would be interesting to show the validation of this gradient, for example, by comparing it to finite differences. Additionally, it would be helpful to demonstrate the convergence rate of the optimization problem as well as the norm of the gradient in both applicative cases.
   Answer: Validation of the gradient, convergence of the numerical schemes and tests on the adjoint operators are available in the git repository of the pherosensor-toobox package that has been the subject of a publication in JOSS (see doi: 10.21105/joss.06863). See e.g. the validation of the numerical schemes here. To highlight the work done on the package, this point has been clarified and a reference to the JOSS publication has been added.

   For the sake of brevity of the article, we did not add plots of the cost function values through the optimization iterations. However, the convergence of the optimization problem as well as the norm of the gradient were monitored when the numerical experiments were performed, especially the cost function values through the optimization iteration were plotted after every numerical experiment. Such figures can be found (for a sighlty more simple test case) in a tutorial available in the git repository or in the package documentation.
   Modifications: see lines 132 and 188

5. **Reviewer comment:** 4. The proposed technique is based on a posteriori approach. It could be useful in applications to periodically update the source predictions based on newly acquired measurements, as done in 4DVAR in meteorology.
   Answer: Yes: we agree with you. In a first approach, we considered the whole time-serie as fully known and focused in how to leverage these data with biology-informed penalties. However, when this type of sensor will be used in an epidemiological context, it would be very informative

to sequentially assimilate data once available.
Modifications: See line 955 where a comment has been added.

6. **Reviewer comment:** 5. How does this approach compare with standard state observers like Kalman filters?
   Answer: As the pheromone propagation model is linear with respect to the source term, if the biology-informed regularization terms are differentiable and quadratic, which is the case e.g. for a population dynamic-informed regularization term with a linear population dynamic model, then they can be expressed as a gain matrix of a Kalman filter and our approach is equivalent to a Kalman smoothing on the whole time window. However, Kalman filtering suffers of two drawbacks that make it less suited in our case.

   First, this equivalence stands if the biology informed-regularization terms are differentiable and quadratic. Thus, it excludes non-differentiable regularization terms such as the LASSO-type regularization terms and non-quadratic regularization terms such as the population dynamic-informed regularization terms with non-linear population dynamics. To tackle this issue, one can use more advance filtering method such as extended Kalman filter or sparse Kalman filter. However, it usually implies that the estimator is no longer optimal or requires more computations.

   Secondly, the basic Kalman filtering technics are derived in a discrete setting, which implies that the different matrices, such as the gain and covariance matrices, should be built and stored, at least partially. Conversely, the variational data assimilation framework we use avoids explicitly building these matrices, reducing the computational requirements.

   On the other hand, the Kalman filtering has the advantage to provide an analysis each time data are available. As mentioned earlier, a more sequential approach will be investigated in a context of sensor repositionning.
   Modifications: See line 960 where a comment has been added.