

# Response to the reviews and relevant changes

B. J. Snow, I. Moulitsas, A. J. Kolios, M. De Dominicis

## RC C2527

### Response to General Comments

1. Page C2527, *“The method used is not described clearly”*

Methodology has been reviewed and enhanced. We have now introduced a new section, namely Section 3 - Methodology.

2. Page C2527, *“What is the advantage of doing this? For example: Medslik II is designed to run on small desktop PCs, but probably could be run even hundreds of times per day, depending on the problem and model set up and CPUs available. Thousands of calculations can be done using CranSLIK (Monte Carlo) but how is this expected to be used in practice?”*

The main advantage of the proposed approach is in the fact that it allows consideration of stochastic inputs in the problem assigning confidence intervals on the measured output quantities. Considering that the problem at hand is governed by highly stochastic variables, the authors strongly believe that this approach becomes pertinent.

Furthermore, the proposed approach is far quicker. As you mention, MED-SLIK II could be run hundreds of times a day. However CranSLIK can run thousands of times a second. This allows for real-time Monte Carlo simulation. At any time the wind and current velocities are known. However their behaviour over the next time period is not. Therefore these values are predicted using the various meteorological models. What CranSLIK allows for is the input of a distribution for the current and wind velocities leading to all possible destinations for the oil spill to be considered. From this, a region can be created in which contains, for example, 95% of all possible destinations. In other words, the suggested approach of employing approximation methods and coupling them with stochastic (i.e. Monte Carlo Simulations) or analytical methods (i.e. First Order Reliability Methods/FORM) reduces computational time significantly, especially when the complexity of systems/cases increases.

We have now introduced Subsection 1.2 - Aims, where we further highlight the novelty of the approach.

3. Pages C2527 - C2528, *“It should be also clearly explained how the performance of the model in this mode could be tested, since it seems it was not possible with the test case here.”*

Performance in aspects of accuracy is assessed through the results of the case study for the reference case. As far as simulation time is concerned, at this stage a comparison cannot be made between direct simulations (running a few thousand times MEDSLIK II deterministically) versus application of the methodology proposed which employs the approximation methods. Such comparison is not feasible as the formed simulations cannot be automated. By definition, the number of deterministic simulations required to train the regression model is substantially smaller than the direct simulations.

4. Page C2528, *“A richer set of test cases should be used to convince the reader of the validity and usefulness of the method. For example, in real transport phenomena, surface patches are often twisted and drawn out into long filaments before being eventually fully mixed. Can this model handle such an event? The ideas presented here have scientific interest and potential application to real problems, but are not adequately or convincingly explained.”*

This paper focuses on the methodology establishment of combining deterministic simulations of MEDSLIK II, approximation methods and probabilistic analysis methods for the probabilistic prediction of oil spills transport and fate under stochastic inputs.

The small scale processes (for example filaments) can not be reproduced using low resolution ocean model (such as MFS) and they are not connected with the oil spill model used (in MEDSLIK II to start with, and therefore in CranSLIK either). CranSLIK, the probabilistic model, can only be as good as the deterministic model used, i.e., MEDSLIK II. Therefore, if a phenomenon is not captured by the deterministic model, then such capability could not possibly be passed to the probabilistic model either. Such clarification is stated in Section 5 - Conclusions.

5. Page C2528, *“A reorganisation of the paper is needed since the Methods are scattered over sections 2, 3, and 4. Section 2 should be drastically reduced and moved to background. It is not adequate as a review, but too long for a summary, especially since it is not the purpose of the paper to try to explain the physics of oil spill fate models.”*

We have performed the reorganisation of the paper. This is also evident by the headings of the sections and subsections and the associated length of text therein.

## Response to Specific Comments

1. Page C2528, *“In the abstract, the reader will probably have no idea what is meant by ”sampled” and ”captured” by ”the Algeria scenario”. In section 1, terms are used but not defined adequately: ”apply sampling”, ”Hypercube”. Otherwise it is an appropriate way to introduce a reader to the topic.”*

**Sampling:** Since the variables are continuous, it is not possible to try all of the values. Therefore samples of each variable must be taken to reduce the continuous variables to sets of discrete values. This is known as sampling, which we consider to be a common term. Several types of sampling are discussed in the paper, the aim being to create a set of discrete values which are representative of the continuous variable.

**Hypercube:** A hypercube is simply an  $n$ -dimensional shape. Here it refers to the design of the simulations. For example, if there were only two variables, the design would be a 2-dimensional hypercube, i.e. a square. Hypercube is a term commonly used in this respect.

The abstract has been reworded to be more clear. In the manuscript we have added further clarifications of the terms we have used. For example, in Section 2 - Uncertainties and Stochastic Modelling, in Subsection 1.2 and 3.2 - Sampling the variables and in Section 4 - Case study where we added the definition for accuracy as it is quantified by the volume of oil explained by the model, i.e. captured, divided by the total volume of oil.

2. Page C2528, *“In section 2, a number of physical and chemical processes are discussed, but it seems these are part of MEDSLIK II, not CranSLIK. This discussion does not seem relevant beyond a single paragraph describing MEDSLIK II capabilities and approach. This should be placed strategically in the new Methods section (see general comments)”*

We have performed the suggested reorganisation of the paper and have created the new Section 3 - Methodology.

3. Page C2528, *“Spreading is described but stated to be very poor approximation, but it is used anyway? There are no references after 1971?”*

Full description of the formulas and all related references on the spreading appear in the cited paper De Dominicis, M., Pinardi, N., Zodiatis, G., and Lardner, R.: MEDSLIK-II, a Lagrangian marine surface oil spill model for short-term forecasting Part 1: Theory, Geosci. Model Dev., 6, 18511869, doi:10.5194/gmd-6-1851-2013, 2013b. We could add more information in this manuscript, but we tend to feel that this would be beyond its aim and scope as we have cited the relevant sources.

4. Page C2528, *“In section 3, there are many confusing aspects. The new Methods section should contain a completely re-written version carefully and logically explaining the technique.”*

The technique is now explained in more detail in the new Section 3 - Methodology.

5. Pages C2528-C2529, *“-Uncertainty: how does it affect the forecast values of wind speed, -How is the hourly value related to the distributions discussed (is a value taken from the forecast model, then a distribution assumed, or are many forecast runs completed as an ensemble?), and -What is meant by apply sampling? line 13 p 7055 -Where do values of wind come from? Which meteorological model? Reference?”*

Quantities such as wind speed is by definition stochastic as they cannot be deterministically predicted, especially considering their time dependence. Due to this presence of uncertainty we refer to forecasting, which is a prediction of the expected value and its distribution rather than absolute calculation, in order to estimate high or low percentiles for deterministic analysis. This new approach takes into consideration a full statistical distribution (of any type) avoiding any subjective decision making. When developing the model a distribution was assumed and then sampled to give a discrete set of velocities. In other words, in CranSLIK an input distribution can be used as an input but how this is obtained is not specific. Assuming a distribution centred on the forecast model is probably the most simple way to do this. As confirmed already, sampling has been defined and further clarified. Finally, in Section 2 - Uncertainties and Stochastic Modelling, we have written that the input is from European Centre for Medium-Range Weather Forecasts (ECMWF).

6. Page C2529, *“In section 4, again there are basic elements that should be put in the context of methodology. For example, Medslik, a deterministic model, is described as the first section under the Probabilistic Assessment. -Is a “solver” a deterministic oil spill fate model? -“Sampling” needs to be explained explicitly.”*

MEDSLIK II was described in the probabilistic assessment because the established deterministic calculations are used as a discrete block on the flow diagram of the probabilistic analysis. However, in the context of the greater reorganisation of the original, the required clarifications have now been provided.

7. Page C2529, *“-Current velocities are generally lower in the Med? This is not common knowledge, if it is even true, so a reference is needed. - Part of the methodology should include the equations used for regression...”*

*what are the relationships between all these component variables: forcing, response, stochastic, deterministic, otherwise it is too vague to follow.”*

In Subsection 3.1 - Choice of Variables, we have provided a detailed context for the wind speed and the distribution used. We cannot think of a neat way of including the equation (polynomial) information in the manuscript. We also do not see why the presentation of the polynomial pertinent to the test case is a valuable addition to the manuscript. Instead, we included this information as a separate file in the repository where the code for the model, along with the test case data and output, can be downloaded from.

8. Page C2529, *“In section 5, the case study is where we can finally hope to see everything working together and the resulting big picture. It is helpful in this regard, but needs more full description of the experiment: forcing, configuration, results, how was the result evaluated and agreement with Medslik and field data quantified.”*

Section 4 - Case Study now presents a more detailed description.

9. Page C2529, *“Section 5.2 is not helpful. Remove or explain more fully.”*

We have kept the section on the sensitivity analysis as this is a standard aspect of the probabilistic modelling.

10. Page C2529, *“Section 5.3 should describe how we can expect to know the distribution of currents, winds, directions in a place and how to fully test this mode against independent data. Figure 7 is not well presented.”*

This has been addressed in the introduced Section 3 - Methodology. We have updated Figure 7 and discussed it in more detail in Section 4.3 - Monte-Carlo simulation.

11. Page C2529, *“Section 5.4 needs to be enriched with some further, detailed discussion of the benefits and limitations of this model: can it produce the results needed, and under what circumstances?”*

CranSLIK can be used to provide a likely region for the oil spill by considering thousands of possible values for wind and current velocities. An approximate distribution is required for this however this can be a distribution centred on the values predicted from meteorological models. This will allow for consideration of errors in the predicted velocities. By considering more advanced approximation methods, we can achieve better accuracy, obviously at the expense of computation time. Discussion has been enriched in Subsection 4.4 - Discussion and Section 5 - Conclusions.

## Response to Technical Corrections

1. Page C2529, “p7049 line 27 ”This” what? (paper, tool, prediction?)”

This tool provides

2. Page C2529, “line 30 ”probable regions for the oil spill”?”

We have now reworded this sentence in Subsection 1.2 - Aims.

3. Page C2530, “p7051 line 2 ”varying models””

We have now deleted this context.

4. Page C2530, “p7052 line 9 ”spreading of oil due to film thickness and area????””

We have now deleted this context.

5. Page C2530, “p7053 line 13 ”can wave water levels?””

We have now deleted this context.

6. Page C2530, “p7056 line 17 ”While the ...acts.” fragment”

We have now corrected this, now in Subsection 1.1 - MEDSLIK II

7. Page C2530, “p7057 line 16-17 ”however ...however.” ??”

We have corrected this, now in Subsection 3.2 - Sampling the Variables.

8. Page C2530, “p7060 and elsewhere what does ”oil captured” mean?”

As mentioned earlier, we have now clarified this.

9. Page C2530, “It is not clear if the simplification of the oil spill fate model is novel”

The novelty of this work lies in the fact that the oil spill transport and fate is assessed through a probabilistic perspective, taking into account the state of the art deterministic tool MEDSLIK II. This is a particularly important addition as the parameters influencing the phenomenon are highly stochastic.

10. Page C2530, “conclusions are weak and not fully justified. Need more convincing discussion and figures.”

We have now expanded this and offered more discussion.

11. Page C2531, *“methodology too vague”*

Methodology has been expanded and clarified through more detailed discussion.

12. Page C2531, *“No mention of related work, is the first such attempt?”*

To the best of our knowledge, this is the first time approximation methods have been used to approximate oil spill fate and transport.

## RC C2992

### Response to General Comments

1. Page C2992, *“The paper presents a simpler model called CranSLIK for oil spill tracking using stochastic methods instead of a Lagrangian model.”*

We would like to clarify that this is not a simpler or simplified model of MEDSLIK II. It is an extension of the method to account for stochastic inputs in the analysis.

2. Page C2992, *“The paper compares the output of the CranSLIK with an open-source Lagrangian model called MEDSLIK II”*

CranSLIK incorporates the deterministic simulations of MEDSLIK II. Any comparison of the efficiency of the two methods is not possible as the approach and purpose is substantially different. The difference of the two approaches has been discussed in Subsection 1.2 - Aims and Section 2 - Uncertainties and stochastic modelling.

3. Page C2992, *“The authors note that the case study is from the Mediterranean sea where the currents are slower. It would also be interesting to see the comparison of results for case studies in other regions.”*

This is just a case study and the fact that lower currents are considered is not a limitation of the method. Limitations of CranSLIK are parallel to those of MEDSLIK II to model the complexity of the physical phenomena. MEDSLIK II has been described in:

De Dominicis, M., Pinardi, N., Zodiatis, G., and Archetti, R.: MEDSLIK-II, a Lagrangian marine surface oil spill model for short-term forecasting Part 2: Numerical simulations and validations, *Geosci. Model Dev.*, 6, 18711888, doi:10.5194/gmd-6-1871-2013, 2013a.

De Dominicis, M., Pinardi, N., Zodiatis, G., and Lardner, R.: MEDSLIK-II, a Lagrangian marine surface oil spill model for short-term forecasting Part 1: Theory, *Geosci. Model Dev.*, 6, 18511869, doi:10.5194/gmd-6-1851-2013, 2013b.

4. Page C2993, *“Many of the ideas are scattered over different parts of the paper; the paper needs to be edited and rearranged to make it more cohesive and focused. The audience of the paper would happen to be more familiar with the physics of oil spills than the stochastic methods used; the paper needs to explain these methods and their advantages vs traditional Lagrangian modeling.”*

In view of both referee’s suggestion, we have performed the reorganisation of the paper. This is also evident by the headings of the sections and subsections and the associated length of text therein. The stochastic approach is now explained in detail in Section 3 - Methodology.

5. Page C2993, *“I think the concepts and ideas presented illustrate the usefulness of a very simplistic model especially in the initial stages of tracking an oil spill. However, the presentation does not make that very clear but it can be achieved with some reorganization of the paper”*

The word simplistic is not representative of the approach. Development of the methodology adopts fully well established non-intrusive approximation methods which allow employment of high fidelity (deterministic) simulation tools such as MEDSLIK II, in order to probabilistically assess the problem with lower computational time and high accuracy. Presentation is now clearer, after the reorganisation that we have performed.

## Response to Specific Comments

1. Page C2993, *“Section 2 be reduced to a summary and moved into the introduction because the objective is not to introduce the physics of oil spill modeling since CranSLIK explicitly ignore them except to chose the variables driving the stochastic model.”*

We have now done this in Section 1 - Introduction and Subsection 1.1 - MEDSLIK II.

2. Page C2993, *“Section 4.1 also be moved to the introduction because it brings focus to MEDSLIK II which is not the new model being developed.”*

We have now done this in Subsection 1.1 - MEDSLIK II.

3. Page C2993, *“Section 3 be expanded, especially increasing the description of uncertainty modeling using Stochastic methods, and also expanding further on the advantages of the Latin-hypercube vs. other methods of sampling.”*

We have now done this in Section 2 - Uncertainties and Stochastic Modelling and Section 3 - Methodology.

4. Page C2993, *“A new methodology section that significantly expands the list at the end of Section 1 should be written. It should also include Sections 4.2 and 4.3 that describes the implementation in much more detail.”*

We have now done this in Section 3 - Methodology.

5. Page C2993, *“Finally, Section 5 needs to be reworked to better explain the significance of the results.”*

We have now done this in Section 4 - Case Study.

6. Page C2993, *“It is also not clear what the criteria for success is.”*

In Section 4 - Case study we added the definition for accuracy as it is quantified by the volume of oil explained by the model, i.e. captured, divided by the total volume of oil.

7. Page C2994, *“Section 5 introduces hindcast modeling without context. The introduction and explanation needs to be expanded.”*

This has been done in Section 4 - Case Study.

8. Page C2994, *“CranSLIK is much simpler than MEDSLIK II so the expectations would be tempered when it comes to accuracy. However, the discussion only points out that CranSLIK is restricted to modeling point spills. The discussion needs to be expanded further. For example, how can CranSLIK be modified to distort the circular shape of the spill. How else can the model be improved: more variables, more or higher order cross-correlations, etc.”*

Non-circular shapes cannot currently be addressed.

9. Page C2994, *“I would also recommend that the regression modelling be described better: order of polynomials, cross-correlations, coefficients, r-squared value, etc.”*

In Subsection 3.1 - Choice of Variables, we have provided a detailed context for the wind speed and the distribution used. We cannot think of a neat way of including the equation (polynomial) information in the manuscript. We also do not see why the presentation of the polynomial pertinent to the test case is a valuable addition to the manuscript. Instead, we included this information as a separate file in the repository where the code for the model, along with the test case data and output, can be downloaded from.

10. Page C2994, “Section 5.3 includes a reference to runs using 5,000 and 10,000 samples. Model results using different sample sizes should be presented to illustrate the significance or lack thereof of sample size.”

We have now run more simulations and included these results.

11. Page C2994, “7049: ”Often the computational cost involved in running a full simulation is too high.” What is the usual run time for Lagrangian oil spill models?”

In Section 3 - Methodology we have now clarified that the deterministic model, MEDSLIK II, requires several minutes per run. 1000 runs using different input parameters would therefore require many hours. This approach avoids this problem by creating a polynomial which maps inputs to a response resulting in 1000 runs being possible in approximately one second.

## Response to Technical Corrections

1. Page C2994, “”the model” or ”the developed model” or variations thereof are used everywhere in the paper. It would be better to use CranSLIK instead. Especially in the results and discussions section where it is often not clear what ”model” refers to.”

We have now made the context more clear.

2. Pages C2994-C2995, “There are repeated sentence clusters, for example, lines 8-11 in 7049 and lines 20-24 in 7050; lines 22-26 in 7051 and lines 4-8 in 7055. They should be rephrased.”

We have now eliminated such repetitions.

3. Page C2995, “There are many run-on sentences that need to be edited. For example: 7050: ”However, significant advances have been made since then, for example, the role of microorganisms in biodegradation is now better understood as discussed in McGenity et al. (2012)” 7056: ”Primarily it is performed to simplify the problem however it also means that the developed methodology can easily be applied to data from any source.” 7057 ”Note however that it is not possible to predict the shape of the resultant graph beforehand however it is expected to be more simple than the test shape.” 7057: ”The result was that the destination can be determined by the current and wind velocities, and the size of the spill depends on the initial spill size as well as the spill age, that is time since initial spill.” 7061 ”It is possible in this case to apply an interpolation since the quantities

*for the next time step are known however this would not be possible in a real scenario.””*

We have now performed these edits.

4. Page C2995, “*General technical correction comment 4: In a few places, some ideas are introduced without prior context. For example, Abstract: What is “the Algeria scenario?” It is introduced without context. 7050: design hypercube is introduced without context.*”

We have rephrased the text in the abstract and further clarified hypercube.