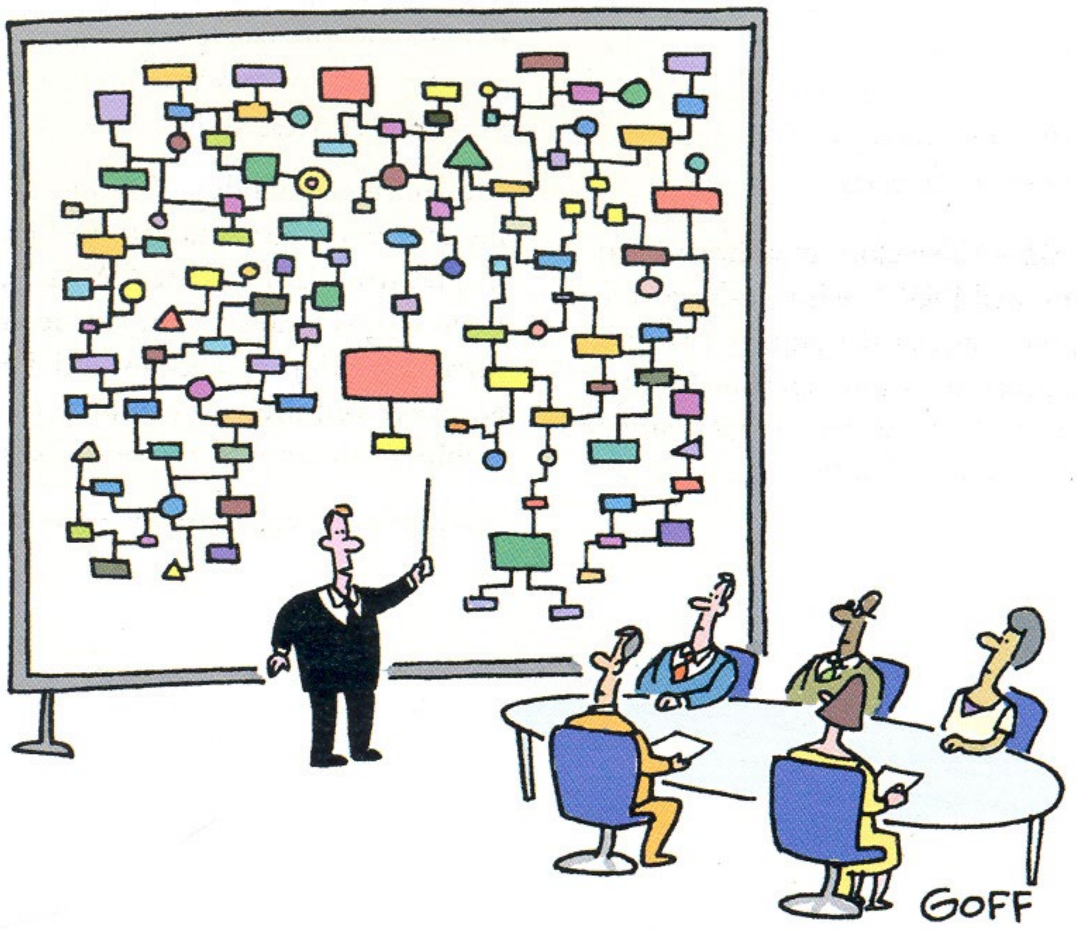
ERA Acute

Pilot Study: Feasibility of uncertainty handling og ERA Acute

Cathrine Stephansen, Akvaplan-niva

Uncertainty handling in ERA Acute



- ERA Acute is a complex, multi-compartment model
  - as compared to "simpler" single compartment models using PEC/PNEC, thresholds etc.
- Complex impact, lag and restoration time functions
- Many parameters
- Uncertainty studies "could go on for ever" if to be accurate...

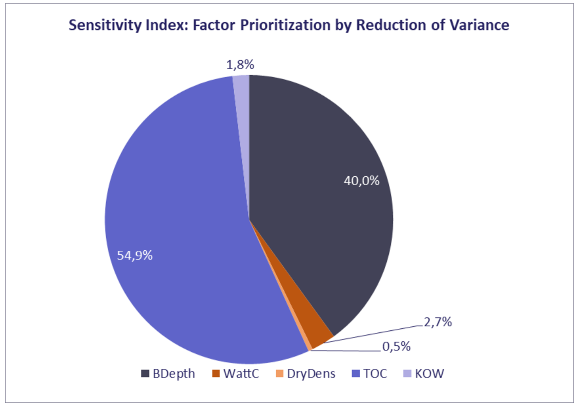
ERAAcute

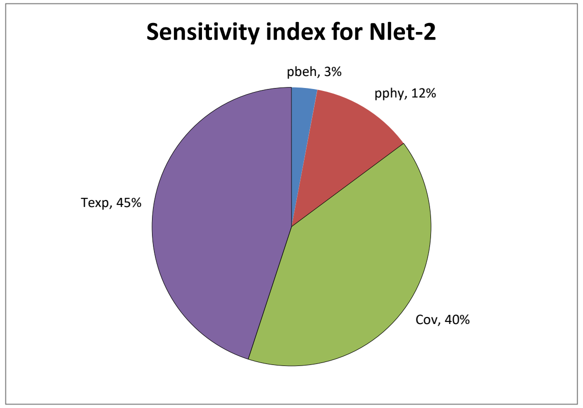
Feasibility study for uncertainty handling

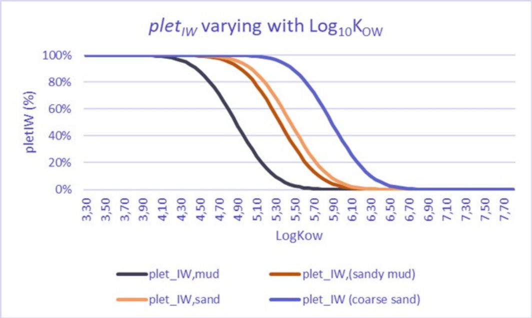
- Goal: *Ensure that ERA Acute does not under-estimate risk*
- Simple solution: Multiply with a “safety factor” to account for uncertainty
  - But what is the correct factor?
  - Many parameters in several functions for each compartment, summarizations from cells/simulations – one safety factor for all, or one for each function?
- All parameters used in the functions have varying degrees of:
  - Natural variability
  - Uncertainty in measurements/assessments/analysis/modelling
  - Importance in the functions with respect to sensitivity

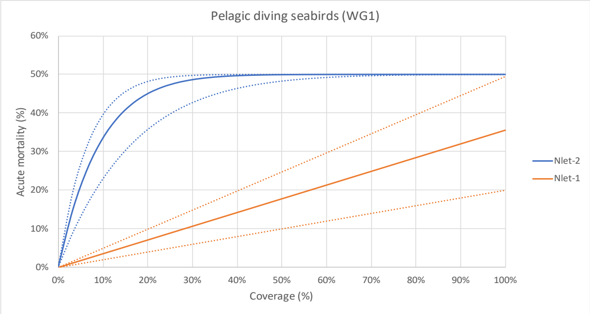
Knowledge from sensitivity studies for ERA Acute

- Phase 4 had carried out sensitivity studies of the functions
    - Extensive sens. testing: Surface and Sea floor
    - Less extensive: Shoreline and water column
    - Deterministic and Stochastic testing
- Had knowledge of which factors the functions are most sensitive to, but:
    - Actual sensitivity results are dependent on the boundaries chosen (stochastic testing results)
    - Boundaries of testing were based on literature values
- Increasing the accuracy of the sensitivity evaluations and finding «exact» uncertainty estimates would require a large project
- Decided for a pragmatic approach possible within a feasibility study.
    - Use results of Phase 4 sensitivity quantification and score using additional qualitative assessment









Examples of results from Phase 4 sensitivity testing of parameters

Scoring system

- Utilize the results of the previous sensitivity testing as much as possible
- Use a scoring system made available by DNV GL developed for input data to MIRA
  - Same scoring still applies for the inputs that are common between MIRA and ERA Acute (inputs to oil drift simulations – driver data)
    - Different oil drift input data for ERA Acute
  - Goal for feasibility study: Score the ERA Acute-specific parameters using the same approach and make recommendations

Scoring of the parameters - questions

- ***Strength of knowledge (function where parameter is used):*** How strong is our confidence in that the risk function in which the parameter is used, is a valid mathematical representation of the mechanism of impact/restoration?
  - Moderate/weak or strong?
  - Based on scientific evaluation of the parameters' place in the function (correctness of functions)
- ***Does the value deviate from the recommended parameter value?***
  - Natural variability of the parameter value entered
    - Do we believe that the values have a high natural tendency to vary from the base case (mean)? E.g. if a (standard deviation) (SD) is quantifiable, this can be used to assess.
  - Uncertainties in the measurements themselves
  - High, moderate or low deviation
- ***Sensitivity of function to parameter (sensitivity index):*** How sensitive is the model/function to variation in this parameter?
  - High, moderate or low
  - From sensitivity testing in Phase 4
- All four compartment parameters were scored

Issues for making recommendations

- Using a safety factor to account for uncertainty deemed unfeasible at this point
  - Finding the right factor not possible at this point
  - Recommended choice of conservative factors amplifies conservativeness throughout summarized risks
  - Adding factors to this could make all activities high risk and reduced decision-making value of ERA Acute
- Probability distributions of each parameter commonly used to illustrate uncertainty (95 % sure the parameter value is 1-100 000…)
  - 49 parameters in total, calculation of values with high and lowestimates for each increases calculation time
  - Deemed unfeasible
- Scoring system gives a qualitative overview of the uncertainty associated with each parameter
  - Important for transparency of the methodology
  - Documentation of the results
- Scoring system gives a ranking of the most important parameters for **improvement of data.**

Sea surface

Function	Main parameter	Strength of knowledge (function where it is used)	Belief that the value may deviate from the average assumption (Natural variation of parameter)	Sensitivity of function to parameter (sensitivity index)	Comments/recommendations on handling to ensure risk is not under-estimated
Impact & Impact time	p_{beh}	Moderate/weak. Due to limited data and large natural variation it is difficult to assign a specific pbeh value. The assumption that behavioural factors will affect p_{exp} is strong.	Moderate	Moderate	A higher value is conservative. Each VEC have three estimates (low, intermediate, high), using high is most conservative. Alternative, use all to obtain larger credible interval.
	Cov	Moderate/weak. The parameter depends on other parameters evaluated as Moderate/weak. The assumption that that exposed area will affect p_{exp} strong.	High	Moderate	A higher value is conservative. Coverage is calculated by the oil drift model. Use Best Practice for ODS set-up to ensure comparable and reliable predictions of the statistic.
	Texp	Moderate/weak. The parameter depends on other parameters evaluated as Moderate/weak. Based on stochastic result (i.e. estimated over the whole simulation period). The assumption that the exposure time will affect p_{exp} is strong.	High	High	A higher value is conservative. Exposure time is calculated by the oil drift model. Use Best Practice input data and setup for the ODS to ensure comparable and reliable predictions
	p_{phy}	Moderate/weak. Due to lack of experimental data, it is difficult to assign a specific p_{phy} values. The assumption that the physiological factors will affect p_{let} is strong	Low/Moderate/High, depending on VEC. Low for seabirds and moderate to high for marina and aquatic mammals and sea turtles	Moderate	A higher value is conservative. Each VEC have three estimates (low, intermediate, high), thus using high is most conservative. Alternative, use all to obtain larger credible interval.
	Th	Moderate/weak. Due to lack of experimental data, it is difficult to assign specific threshold levels for lethal oil film thickness	Moderate	High	A threshold value, lower value is conservative. Oil thickness is calculated by the oil drift model. Use Best Practice for ODS to ensure comparable and reliable predictions. Based on present knowledge, reducing Th from 10 mm to 2 mm, increases the impact with a factor of approximately 2.0-2.5, depending on the distribution of the VEC and the distance to the release point.
	N per cell	Moderate/weak. Depends on the quality of the data received from the data provider. The quality of the data for the NCS is considered high.	High	Moderate/high	Use the best available data to reduce uncertainty. Use the same data for comparable studies. The definition of a “population” is important.
Lag-time	Nhab	Moderate/weak. The function includes various not well-defined or understood subtle effect other than acute mortality.	High	Moderate	Using the function will increase the total recovery time, typically with 5-30% of the shoreline lag-times but depending on the importance of the affected shoreline habitats.
	SF	Moderate/weak	High	Moderate	
	t_lag (shoreline)	Moderate/weak. Due to lack of experience data, it is challenging to assign specific lag time periods for different types of shoreline habitats.	High	High/Moderate	Higher values are more conservative. Standard values for SF for different VECS and/or area are not derived. May use the same data as for calculating acute mortality (filtered for shoreline cells).
Restoration time	R	Moderate/weak. The logistic discrete population model is a simplification of real-world population dynamics. Common R values are used for different species and populations as a standard (see b)	Moderate/high	High	Lower values are more conservative. The R values are conservative compared to the damage keys used in MIRA (using standard values for b, K and TLR). Field validation studies indicates that the model performs reasonably well, for population not inhibited by unknown extrinsic factors (using standard R, b, K and TLR values).
	b	Moderate/weak. The parameter determines the strength of intraspecific competition; a simplification of real-world population dynamics.	High	High	Lower values are more conservative. Used to reflect population growth in population inhibited by unknown extrinsic factors or the general status of the population ("poor", "intermediate", "good"). Use low b values to further increase the conservatism of the population model predictions.
	K	Moderate/weak	High. Large fluctuations of population size above and below carrying capacity is common in nature	High	The carrying capacity of the environment (K) is the maximum population size that the environment can sustain. It is set equal to the population size before the oil spill release (100%) and is used as a reference point for when the population is considered recovered.
	TRL	Moderate/weak Cut off to avoid tres = ∞ in a logistical growth model.	High	High/moderate for t_{res}, Moderate/low for RDF (effect varies with percentage population loss)	Higher values are more conservative. Can be chosen differently for higher level of conservatism. Using values above 95% may lead to unrealistic long Restoration times

Surface summary - prioritised

Function	Parameter	Recommendation for improvement
Impact & Impact time	Cov	Use oil drift model that uses a state-of-the art calculation of oil coverage above the threshold on the surface with best practice settings
	Texp	Use oil drift model that uses a state-of-the art calculation of the time with oil above the threshold level on the surface, with best practice settings
Restoration time	R	Net fundamental growth rate can be inherently different for different populations. Updating the knowledge and adjusting the growth rate/using individual growth rate for populations at risk would improve certainty.
	TLR	Current restoration function is asymptotic, the threshold level for when the population is recovered is highly sensitive
	b	Use lower values for conservativity, and best practice recommendation

Water column

Function	Main parameter	Strength of knowledge (function where it is used)	Belief that the value may deviate from the average assumption (Natural variation of parameter)	Sensitivity of function to parameter (sensitivity index)	Comments/recommendations on handling to ensure risk is not under-estimated
Impact	Plet, THC Extracted from SSD-curve	Strong	Moderate. SSD-curve based on LC50 for 24 species	High	Estimated from THC and a log-normal SSD curve with standard deviation of 0.32. A lower standard deviation is conservative (shift the SSD curve to higher THC values).
	THC	Moderate/weak. Vertical maxima, THC includes numerous components with varying toxicity	High	High	THC is calculated by the oil drift model. Use Best Practice for ODS set-up to ensure comparable results. Use a concentration grid (with many layers) that cover the same water column where the fish egg/larva are distributed
	Pfrac (SD and species sensitivity)	Strong. Estimated in OSCAR during the ODS using a simplified QSAR method	Moderate	High /Moderate/ Low (depending on setting)	Estimated by OSCAR during the ODS (potential acute mortality in a cell). Standard deviation (SD) of the SSD and the species sensitivity may be adjusted before one run the ODS. The species sensitivity is a safety factor. The OSCAR database LC50 values will be divided by this factor, accounting for more (factor > 1) or less (factor < 1) sensitive fish larva/egg.
	N per cell	Strong. Depends on the quality of the data received from the data provider. Compared to e.g. birds the distribution is to a large degree dependent on sea currents	Moderate	Moderate/high	Use the best available data to reduce uncertainty and increase the quality of the predictions. Use the same data for comparable studies.
Recovery	CritDens%	Moderate/weak	High	High (threshold level between two methods with different conservatism)	Higher values are more conservative Expresses the threshold for when a direct relationship is modelled between larval mortality and recruitment reduction
	CritOilMort (%)	Moderate/weak	High	High (threshold level between two methods with different conservatism)	Lower values are more conservative. Expresses the threshold mortality of eggs and larvae for which a proportionate relationship is calculated between killed larvae and reduced recruitment
	Annual natural mortality of immatures (%)	Moderate/weak	Moderate/high	Not tested	
	Annual natural mortality of matures (%)	Moderate/weak	Moderate/high	Not tested	
	Age at recruitment (year)	Moderate/weak	Low /moderate	Not tested	
	Age at first spawning (year)	Moderate/weak	Low	Not tested	
	Maximum age (year)	Moderate/weak	Low	Not tested	

Water column summary - prioritised

- Toxicity functions as per current science, less uncertainty than is compartments with no experimental data

Function	Parameter	Recommendation for improvement
Impact & Impact time	p_{let}	Using an oil drift model that uses a time-step-calculated QSAR-approach to calculated fraction killed (p_{le}) considered most valid approach
Restoration time	CM	Use a best practice recommendation for setting the Critical Mortality value for when the gate model is used

Shoreline

Equation	Main parameter	Strength of knowledge (function where it is used)	Belief that the value may deviate from the average assumption (Natural variation of parameter)	Sensitivity of function to parameter (sensitivity index)	Comments/recommendations on handling to ensure risk is not under-estimated
Impact	Tidal range (m)	Moderate/low	Moderate/low. (coastal tidal ranges vary considerably depending on the volume of water adjacent to the coast, and the geography of the basin. Tidal range also varies depending on the locations of the moon and sun).	Low	Lower values are more conservative. The parameter is cell specific and is used to estimate oil thickness.
	Slope (°)	Moderate/low	High/moderate. (the slope of the beach may vary considerable with a shoreline habitat type)	High	Higher values are more conservative. The parameter is ESI specific and is used to estimate oil thickness.
	OHC	Moderate/low	High/moderate. (the distribution of oil along the shoreline will also depend on factors such as current, wind, geography, that are difficult to accurate estimate outside the oil drift model)	Moderate/high	Higher values are more conservative. The parameter is ESI specific and is used to distribute the stranded oil mass along the shoreline in a cell. Higher value means that more of the stranded mass is allocated to the shoreline habitat
	Patchiness factor	Moderate/low. Due to lack of experience data, it is challenging to assign a specific patchiness factor	High. Patchiness of oil may range from 1-100%	High	Lower values are more conservative. Fixed look-up values
	Th	Moderate/low. It is difficult to assign a specific threshold level for lethal oil film thickness for invertebrates and vegetation	Moderate	High (threshold value)	Higher values are more conservative. Threshold level for impact, 0.1 mm for invertebrates and 1.0 mm for wetland vegetation
	Stranded mass (ton)	Moderate/low. Basis for calculating film thickness	High	High/moderate(proportional)	Higher values are more conservative. Stranded mass is calculated by the oil drift model. Use Best Practice for ODS to ensure comparable and reliable predictions.
	Shoreline length (km)	Strong. Depends on the quality of the data received from the data provider.	Low/moderate	High (proportional)	Use the best available data to reduce uncertainty and increase the quality of the predictions. Use the same data for comparable studies.
	Shoreline rankings	Strong. Depends on the quality of the data received from the data provider.	Moderate	High for recovery (lag-time and restitution)	ESI rankings; 1 least sensitive, 10 most sensitive
Lag-time	Lag-time	Moderate/low. Due to lack of experience data, it is challenging to assign specific lag-time periods for shorelines	High/moderate. Variable and to a large degree depending on weather conditions	High	Fixed look-up values
Recovery	Recovery	Moderate/low. Due to lack of experience data, it is challenging to assign specific restitution time periods for shorelines	High Variable depending on invertebrate and flora communities	High	Fixed look-up values

Shoreline summary - prioritised

Compartment	Parameter	Recommendation for improvement
Impact & Impact time	Mass	High importance but proportional. Use oil drift model that uses a state-of-the art calculation of beached mass, with best practice settings,
	Patchiness factor	The value is a fixed value based on research. Lack of data available, could be improved with more research
	Slope	ESI-specific. Use best practice ESI dataset.
Lag-time/ Restoration time	Lag-time/ Recovery time	Fixed values that could be improved with more research

Sea floor

Main parameter	Strength of knowledge (function where it is used)	Belief that the value may deviate from the average assumption (Natural variation of parameter)	Sensitivity of function to parameter (sensitivity index)	Comments/recommendations on handling to ensure risk is not under-estimated
Mixing depth	Strong/ moderate. Knowledge of what constitutes the bioturbation depth is relatively strong	High uncertainty	40.0 % HIGH	A lower value is conservative, lower values are default for all substrates based on size of typical burrowing fauna in substrate. High natural variation: Either look for local real values or use conservative value.
Dry density	Strong	Low	0.5 % LOW	Marine Geochemistry gives general values. Low sensitivity, use defaults.
Water Content	Strong	Low/Moderate	2.7 % LOW	Use lower values as conservative.
Total org. Carbon	Strong (EqP accepted methodology)	High	54.9 % HIGH	Use conservative (lower) values. Lower values lead to higher toxicity and shorter restoration times. (Higher TOC sequesters THC in sed.)
KOW	Strong (EqP accepted methodology)	Moderate	1.8 % LOW	Value calculated based on typical components with affinity to organic carbon in sediment. Use as implemented, can be changed, but has low impact on result.
Plet (SSD-curve used)	Strong	High to low depending on species sensitivity	High	Conservativity already implemented by the curve being conservatively extrapolated from the LD5-value from a large and QA'ed set of data (Nilsen et al. 2006). SSD curves are accepted methodology and inherent safety factor used.
THCsed (used as input from OSCAR	Strong knowledge of place in ERA Acute function	Is calculated by the OD model. SD is low within calculations in same model, may vary a lot between inputs from different models	High (proportional)	THCsed calculations in OSCAR do currently not take into consideration the grain size or TOC-content of the substrate (these factors are used by ERA Acute to modify the exposure in the initial calculations. No conservativity is included, but the other factors are chosen conservatively. The calculations in sediemnt in OSCAR are undergoing improvements, e.g. by possible inclusion of marine snow.
THC (WC)	Strong knowledge of place in SSD-curve	High uncertainty and the THC concentration is a time-averaged concentration	High (proportional)	The concentration is calculated as a time-averaged THC-value. This is a weakness in the approach. Use of dynamic time-steps output options (e.g. proposed in the ERA Acute Dynamic Risk Assesment incl. MIZ-proposal) could improve this. HOWEVER Conservativity is applied as we currently do not have available from OSCAR the THC-conc. in the lower WC, and therefore use the upper layers as for compartment WC. This is conservative.
N	High strength of knowledge	Moderate	High (proportional	Use quality data on presence or habitat area/fractions. Sampling of benthic species may lead to uncertainties, use data that are based on accepted sampling methods by accredited data sources.
$C_{threshold,sed}$	Moderate strength of knowledge of function	High	High	Concentration of THC at which effects on faunal communities in sediment cannot be detected in monitoring studies (Renaud et al. 2008). Species may be more sensitive or less.
$C_{benchmark-max,sed}$	Moderate strength of knowledge of function	High	High	Value representing the maximum value at equilibrium. Based on data from the MOD data base (North Sea).
20 years def value	Moderate strength of knowledge of function	High	High	The based on MOD data from North Sea, sandy bottom, few sites have data on restoration times after use of oil based drilling muds.
SF	Moderate strength of knowledge of function.	High	High (proportional)	Theoretical calculation of the leaching of THC from organic carbon, simplified approach based on physical-chemical properties of THC bound to organic carbon in sediments (resuspension and redistribution may vary between substrates and is not included). The SF was introduced to the function to modify the calculated restoration time
Tlag (HARD)	Fixed value	High	High	Very little research available after oil spills affecting deep sea corals. Comparable incident DWH not yet restored.
Tres (Hard)	Fixed value	High	High	Very little research available after oil spills affecting deep sea corals. Comparable incident DWH not yet restored.

Sea floor summary - prioritised

Compartment	Parameter	Recommendation for improvement
Impact	TOC	Total Organic content in the soft substrate determines the partitioning between oil adhered to the substrate and oil that is bioavailable in interstitial or gut water, and thereby the exposure and lethality. The value may vary a lot regionally depending on the background concentration of organic matter and substrate type. Monitoring studies could include this parameter for regionally/nationally improved quality of the substrate data
	BDepth	Mixing depth scales the result proportionally and varies with the type of burrowing fauna. The variation in results from different studies is high. Monitoring studies could include this parameter for regionally/nationally improved quality of the substrate data
	WC oil concentration	Exposure through water column determines much of the impact for all feeding modes that have exposure though water column. Best result if using oil drift modelling that provides a separate water column concentration from the bottom layer.
Impact, Lag- and Restoration	THCsed	Start-value of oil concentration in the soft substrates. Use an oil drift model that provides a state-of-the-art calculation of oil in the sediment corrected for the substrate type (TOC-content).

Recommendations for ERA Acute Guideline

1. Use the reasonably *conservative* parameter values included in the method reports and current guideline
   - Use the conservative "QSAR" approach to estimate larvae losses in water column impact calculations, not THC-time-weighted average, more in line with SYMBIOSES results
2. Continuous improvement of parameter certainty
   - Use the scoring system to identify and prioritize parameters for which more accurate values would reduce uncertainty (knowledge/accuracy value for research money)
   - When researching new data:
     - Use *quality* data sources from acclaimed institutions
     - Seek *improved* data for the factors to which the model is most sensitive to where possible
3. Develop regional Industry standards (i.e. National)
   - Reduces variability between analyses (all "equally wrong"…) within the region
   - Data sets, input data, best practices, common recommendations for parameter values etc. should be common standard
   - Change standard parameter after consensus and QA when new developments are available

Thank you!

The ERA Acute project is carried out by a consortium of industry partners (Statoil, Total, Norwegian Oil and Gas Association) and experts in environmental risk analysis (Acona, Akvaplan-niva (project manager), DNV-GL and SINTEF), supported by the Research Council of Norway.