

# Implementation of OMV Machine-Learning augmented workflows to support scenario evaluations under uncertainty in DELFI

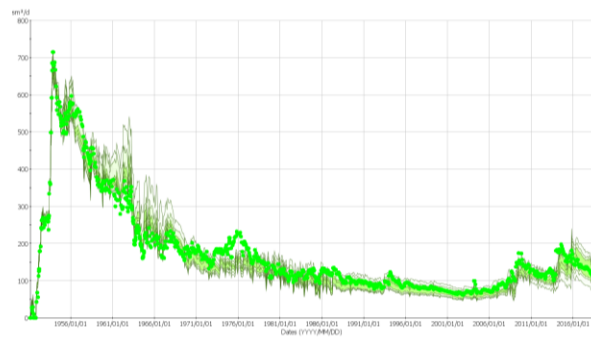
Gisela Vanegas Cabas, OMV  
David Böck, OMV  
Torsten Clemens, OMV  
Maria Chiotoroiu, OMV  
Mathias Bayerl, OMV  
Adrian Anton, SLB  
Andrew Round, SLB

Lucerne, September 2022

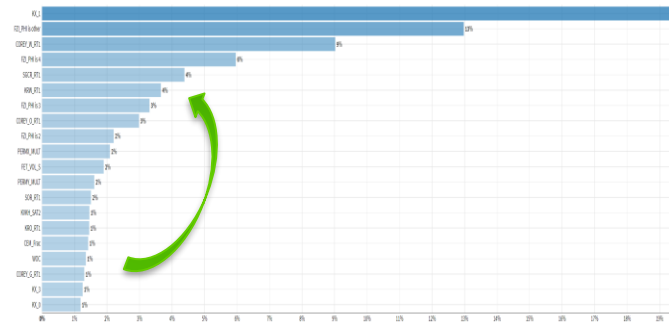


OMV Exploration & Production

# The challenge: scenario evaluations under uncertainty



Numerical models can be conditioned to many different types of measured data, this is an ill-posed problem, a multitude of different parameter combinations can lead to acceptable agreement with the observed data.



The final model ensemble conditioned to all observed data can be used to forecast under uncertainty where changing production conditions might change non-sensitive parameters in the history to sensitive parameters in the forecast.



Optimizing field re-developments leads to increasing value ranging from 5 to 50M EUR per field. The risk of failure can be reduced from 40% to 10% by using probabilistic workflows.

# OMV Machine-Learning augmented workflows

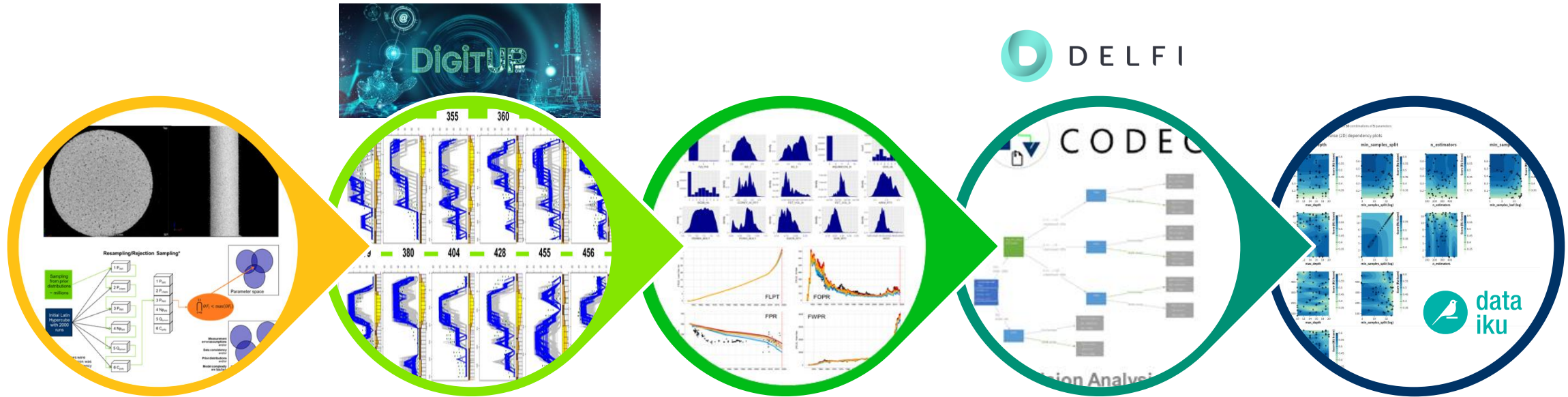
Instead of finding the best parameter combination by minimizing the mismatch, prior distributions are used for the different uncertain parameters which are then conditioned to the various observed data obeying Bayes' updating rule.

The Random Forest Machine Learning algorithm is used to generate a large variety of parameter combinations in an acceptable agreement with the various sources of observed data, addressing also parameter interactions.

Machine Learning Workflow 2020



# OMV Machine-Learning Workflows Digital Journey



2019: Concept from coreflow simulation study  
SPE-200578-MS

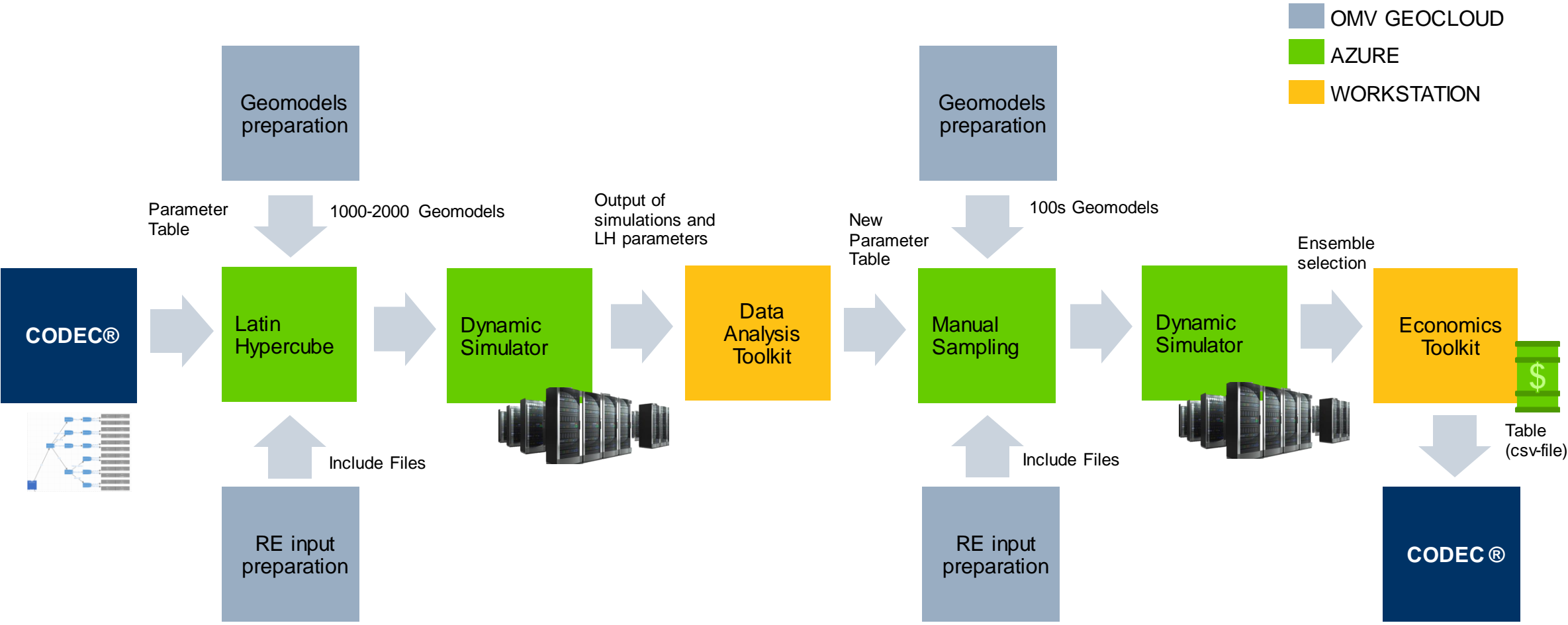
2020 Q1: Method applied for water saturation initialization  
SPE-203384-PA

2020 Q2-Q4: First workflow application for field studies  
SPE-205188-PA &  
SPE-208194-MS

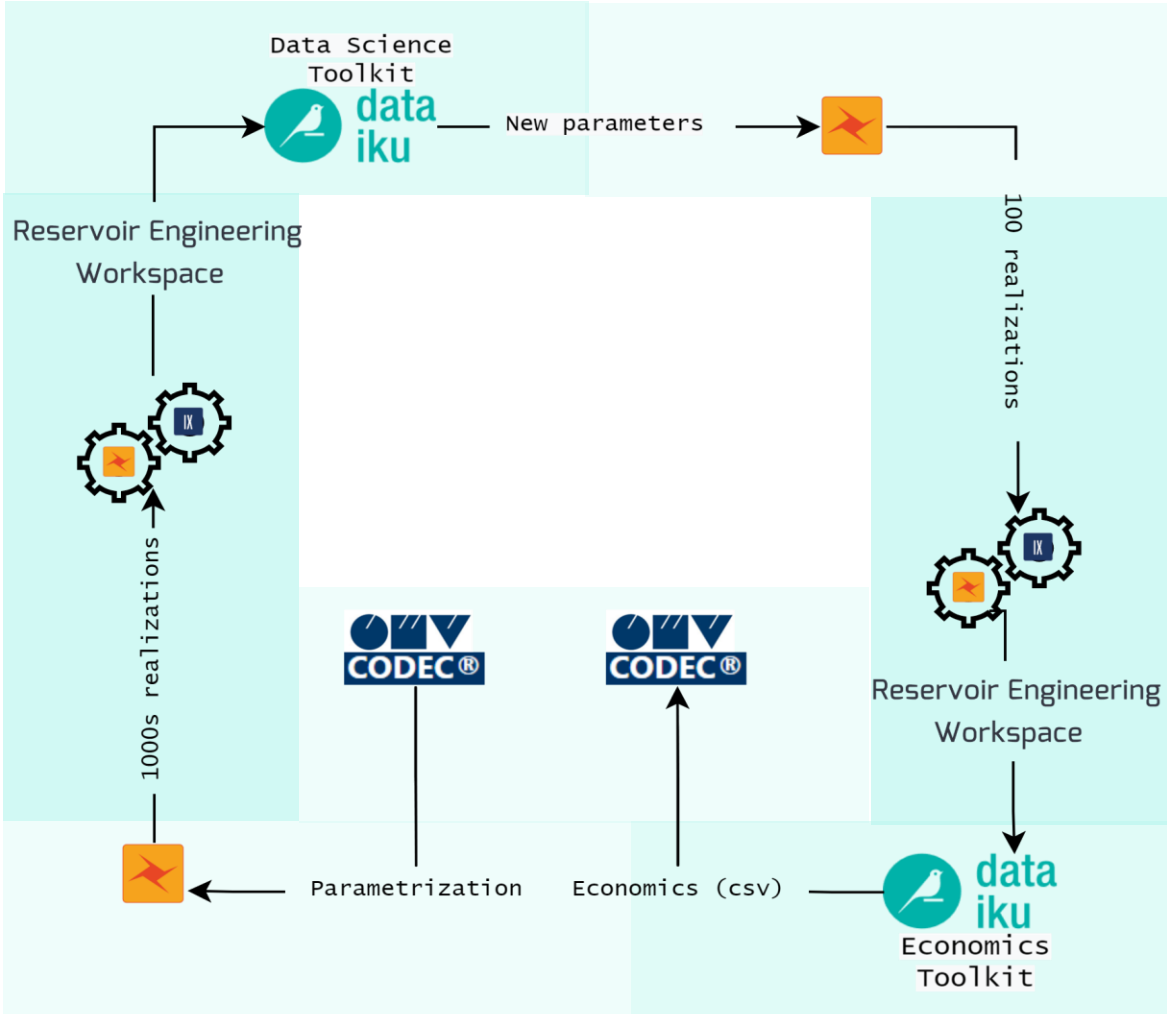
2021: First Co-Development of Stochastic Modelling in Delfi: Clustering & CODEC®







2022: Co-Development of Stochastic Modelling in Delfi: **ML Workflows**

# Original workflow

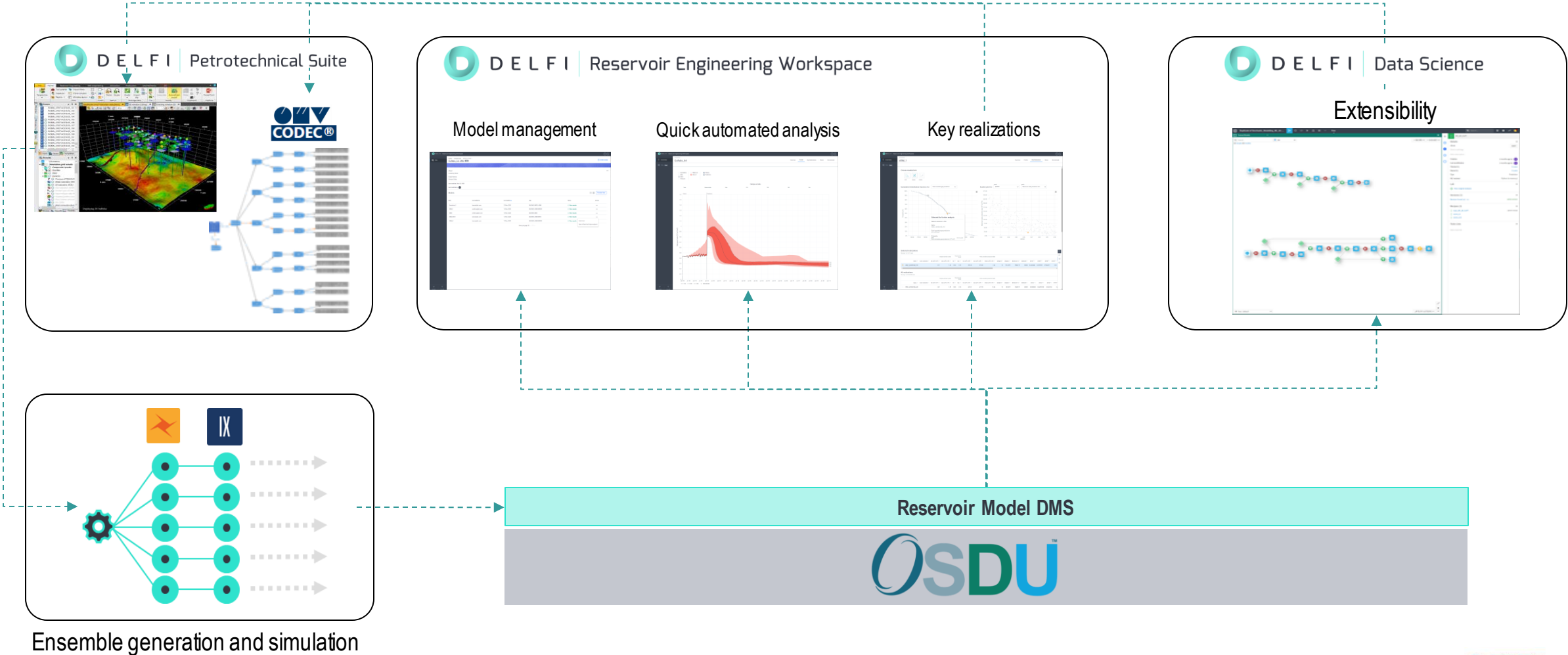


# Simplified workflow

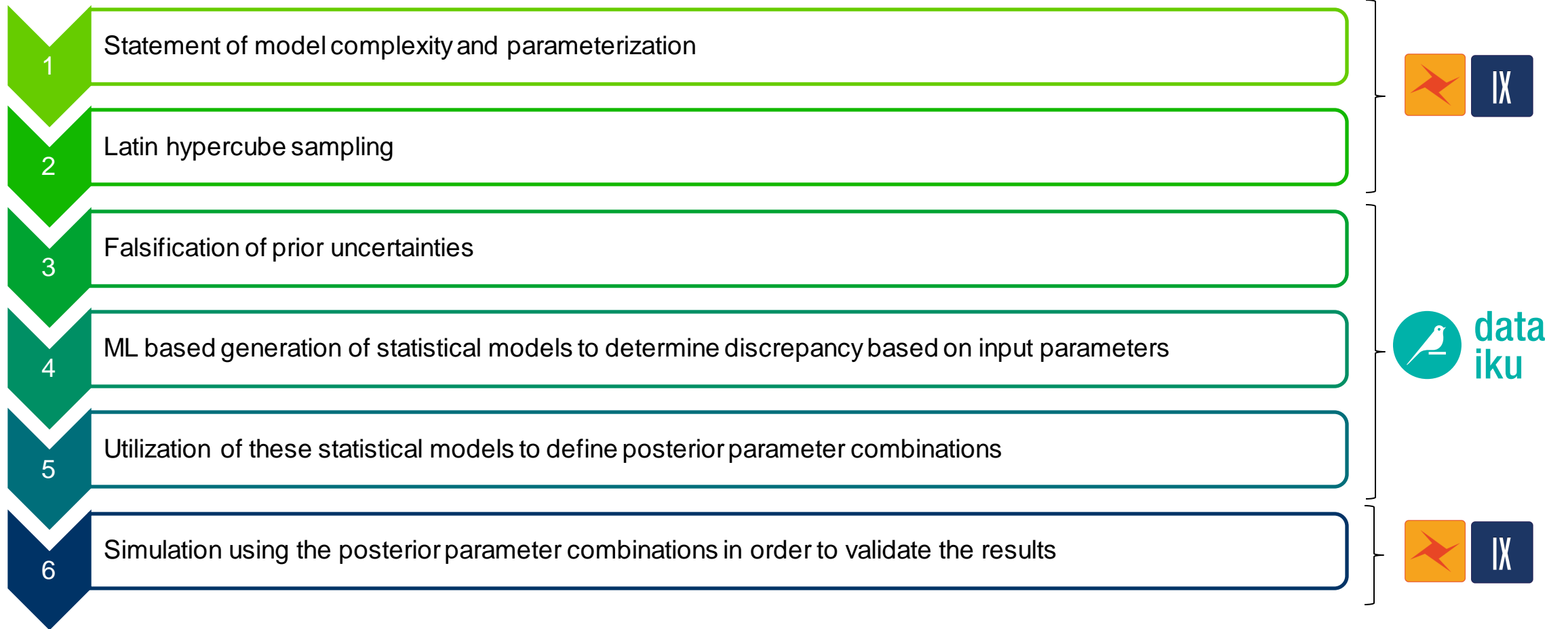


-   DELFI | Petrotechnical Suite
-   DELFI | Reservoir Engineering Workspace
-   DELFI | Data Science

# Stochastic Modelling Architecture



# OMV ML Model Conditioning Approach

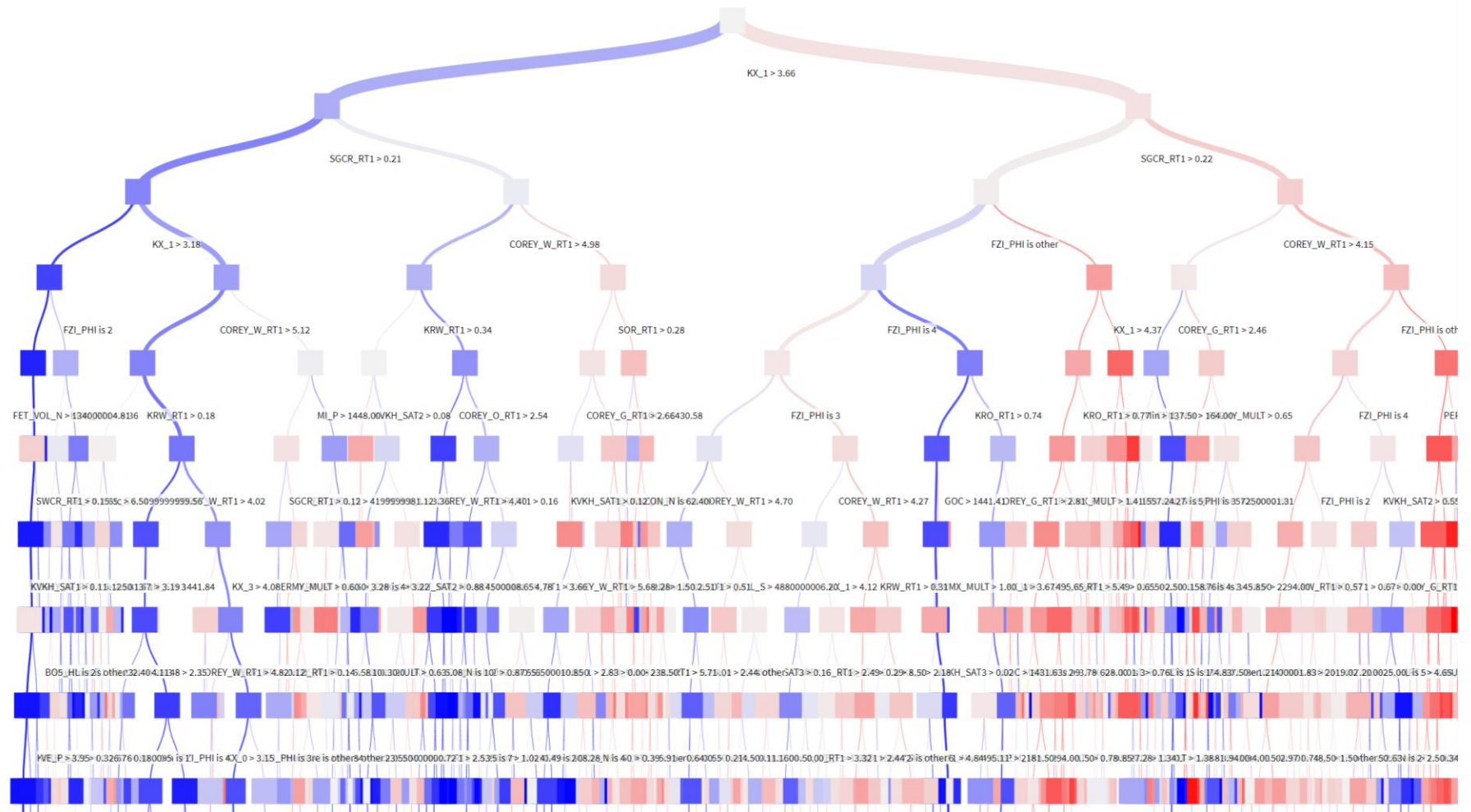




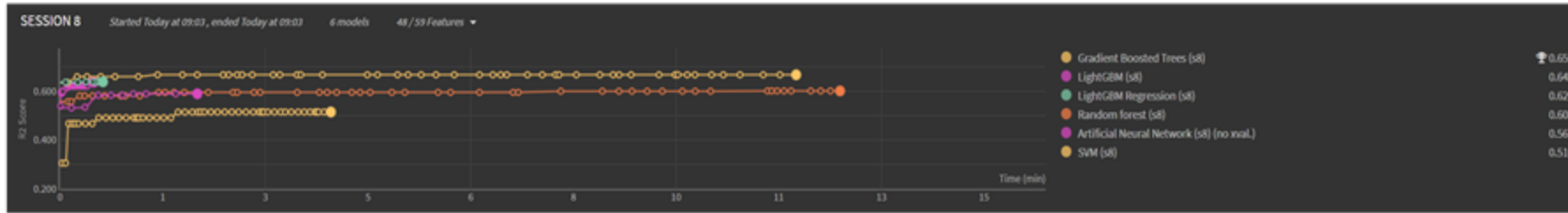
# Statistical models: Random Forest

The methodology allows

- Categorical, continuous geological and dynamic model parameters
- Calibration to multiple data at well and reservoir level
- Decision trees created for each observed data type (objective functions) based on the importance of variables in the regression.



# Statistical models: Exploring alternative algorithms

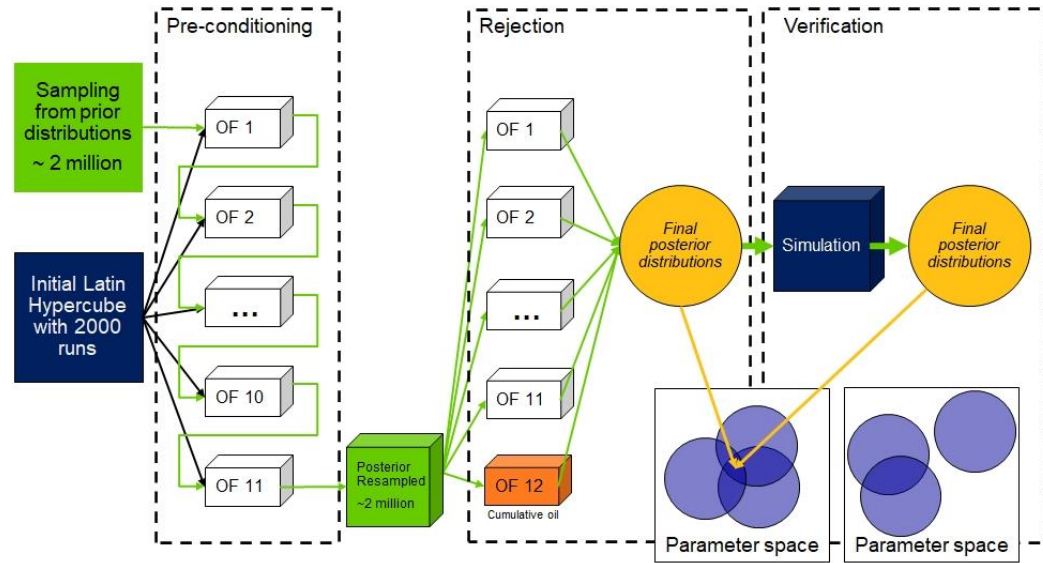


DELFI Data Science

- Easy access and testing capabilities to various ML algorithms
- Compare different regression models, evaluate impact, and adapt the workflows



# Utilization of statistical models: from concept to scripts

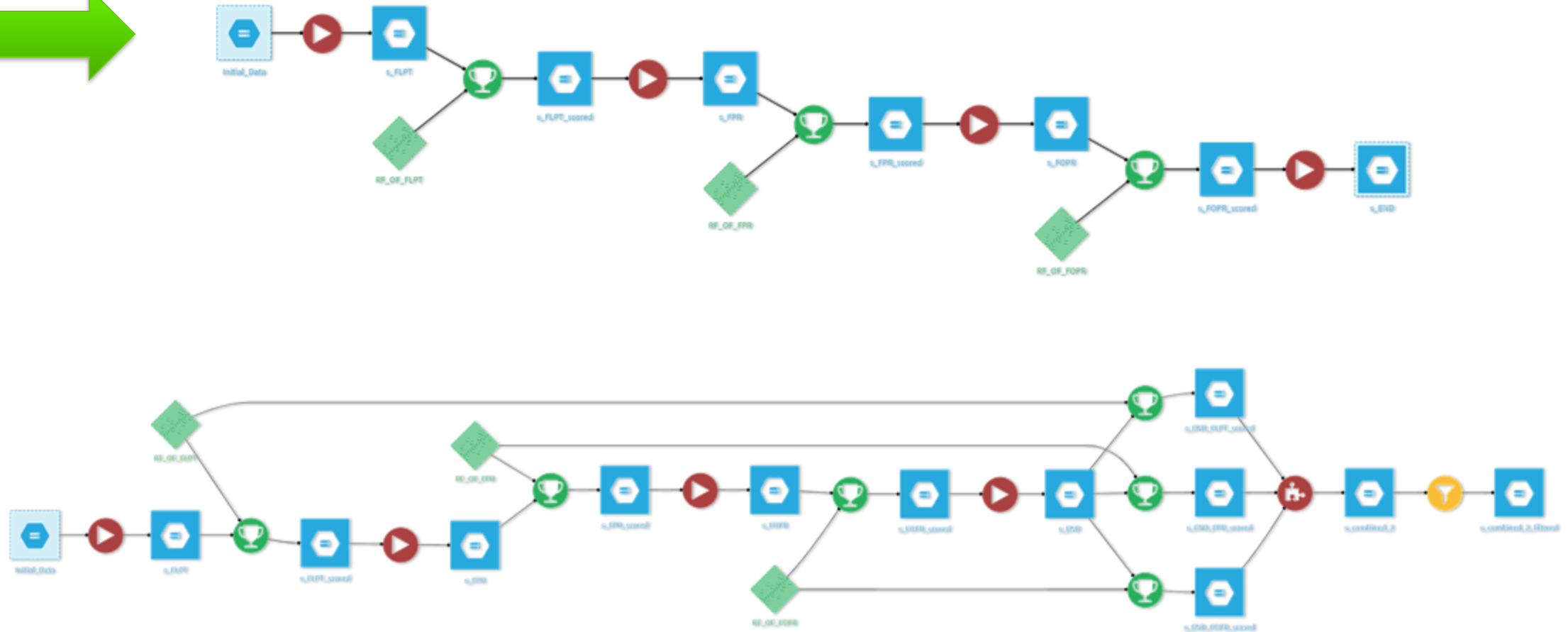
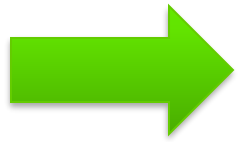


```

-----
# RandomForest Analysis
-----
508
509
510 # Initialize the first matrix to add on
511 analysis_RF <- RF_cl[[1]]$importance
512 print(analysis_RF)
513
514 # Build a add sequentially the matrices of importance from the RandomForest RF_cl
515
516 for(i in @int_OF){
517
518   analysis_temp <- RF_cl[[i]]$importance
519
520   #merge the previous matrix (the initial one) and the temporary one
521   analysis_RF = merge(analysis_RF, analysis_temp, by = "row.names", all=TRUE)
522   # replaces the NA "not a number" or "no data value" by 0
523   analysis_RF[is.na(analysis_RF)] <- 0
524
525   #reshapes the matrix to keep the row names
526   newmatrix = as.matrix(analysis_RF[-1])
527   rownames(newmatrix) <- analysis_RF[-1]
528   analysis_RF = newmatrix
529
530   #sum the columns of the matrix per rows
531   analysis_RF = apply(analysis_RF, 1,sum)
532
533   # the final matrix contains the variables used in the RandomForest and the sum of their "Increase Node Purity"
534   cat("Final list of variables and corresponding sum of nodes purity is \n")
535   print(analysis_RF)
536   #Sort and print the matrix in the order of importance
537   analysis_RF_results=sort(analysis_RF, decreasing = TRUE)
538   print(analysis_RF_results)
539
540
541 ##### Count the occurrences
542 analysis_occ <- RF_cl[[1]]$importance
543 #transform the importance into 1
544 analysis_occ[analysis_occ > 1] <- 1
545
546 for(i in @int_OF){
547
548   analysis_temp <- RF_cl[[i]]$importance
549   #transform the importance into 1
550   analysis_temp[analysis_temp > 1] <- 1
551
552   #merge the previous matrix (the initial one) and the temporary one
553   analysis_occ = merge(analysis_occ, analysis_temp, by = "row.names", all=TRUE)
554   # replaces the NA "not a number" or "no data value" by 0
555   analysis_occ[is.na(analysis_occ)] <- 0
556
557   #reshapes the matrix to keep the row names
558   newmatrix = as.matrix(analysis_occ[-1])
559   rownames(newmatrix) <- analysis_occ[-1]
560   analysis_occ = newmatrix
561
562   #sum the columns of the matrix per rows
563   analysis_occ = apply(analysis_occ, 1,sum)
564
565   # the final matrix contains the variables used in the RandomForest and their number of time they occurred
566   cat("Final list of variables and corresponding sum of nodes purity is \n")
567   print(analysis_occ)
568   #Sort and print the matrix in the order of occurrences
569
570 }
571
572 }
573
574 }
575
576 }
577
578 }
579
580 }
581
582 }
583
584 }
585
586 }
587
588 }
589
590 }
591
592 }
593
594 }
595
596 }
597
598 }
599
600 }
601
602 }
603
604 }
605
606 }
607
608 }
609
610 }
611
612 }
613
614 }
615
616 }
617
618 }
619
620 }
621
622 }
623
624 }
625
626 }
627
628 }
629
630 }
631
632 }
633
634 }
635
636 }
637
638 }
639
640 }
641
642 }
643
644 }
645
646 }
647
648 }
649
650 }
651
652 }
653
654 }
655
656 }
657
658 }
659
660 }
661
662 }
663
664 }
665
666 }
667
668 }
669
670 }
671
672 }
673
674 }
675
676 }
677
678 }
679
680 }
681
682 }
683
684 }
685
686 }
687
688 }
689
690 }
691
692 }
693
694 }
695
696 }
697
698 }
699
700 }
701
702 }
703
704 }
705
706 }
707
708 }
709
710 }
711
712 }
713
714 }
715
716 }
717
718 }
719
720 }
721
722 }
723
724 }
725
726 }
727
728 }
729
730 }
731
732 }
733
734 }
735
736 }
737
738 }
739
740 }
741
742 }
743
744 }
745
746 }
747
748 }
749
750 }
751
752 }
753
754 }
755
756 }
757
758 }
759
760 }
761
762 }
763
764 }
765
766 }
767
768 }
769
770 }
771
772 }
773
774 }
775
776 }
777
778 }
779
780 }
781
782 }
783
784 }
785
786 }
787
788 }
789
790 }
791
792 }
793
794 }
795
796 }
797
798 }
799
800 }
801
802 }
803
804 }
805
806 }
807
808 }
809
810 }
811
812 }
813
814 }
815
816 }
817
818 }
819
820 }
821
822 }
823
824 }
825
826 }
827
828 }
829
830 }
831
832 }
833
834 }
835
836 }
837
838 }
839
840 }
841
842 }
843
844 }
845
846 }
847
848 }
849
850 }
851
852 }
853
854 }
855
856 }
857
858 }
859
860 }
861
862 }
863
864 }
865
866 }
867
868 }
869
870 }
871
872 }
873
874 }
875
876 }
877
878 }
879
880 }
881
882 }
883
884 }
885
886 }
887
888 }
889
890 }
891
892 }
893
894 }
895
896 }
897
898 }
899
900 }
901
902 }
903
904 }
905
906 }
907
908 }
909
910 }
911
912 }
913
914 }
915
916 }
917
918 }
919
920 }
921
922 }
923
924 }
925
926 }
927
928 }
929
930 }
931
932 }
933
934 }
935
936 }
937
938 }
939
940 }
941
942 }
943
944 }
945
946 }
947
948 }
949
950 }
951
952 }
953
954 }
955
956 }
957
958 }
959
960 }
961
962 }
963
964 }
965
966 }
967
968 }
969
970 }
971
972 }
973
974 }
975
976 }
977
978 }
979
980 }
981
982 }
983
984 }
985
986 }
987
988 }
989
990 }
991
992 }
993
994 }
995
996 }
997
998 }
999
1000 }
    
```

SPE-205188-PA • Conditioning Model Ensembles to Various Observed Data (Field and Regional Level) by Applying Machine Learning Augmented Workflows

# Utilization of statistical models: to code-free dynamic workflows



# Output and Validation

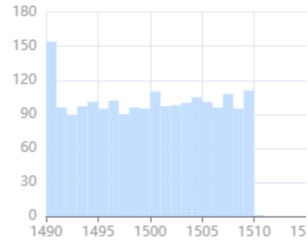
The use of statistical ML models ensures that non-sensitive parameter ranges are kept aligned to the prior distributions

Fast feed-back on parameter distribution and validity of ensembles

Prior Distributions

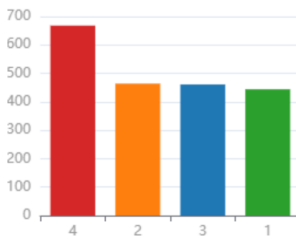
# WOC

Histogram



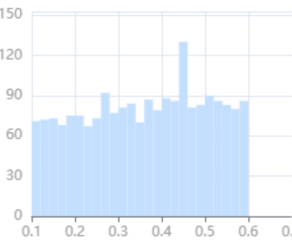
▲ FZI\_PHI

Histogram



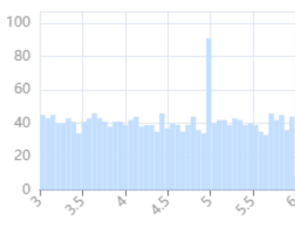
# KRW\_RT1

Histogram



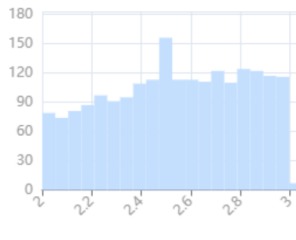
# COREY\_W\_RT1

Histogram



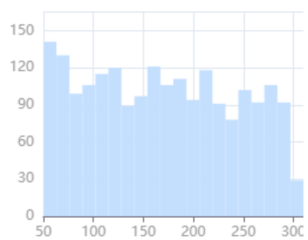
# COREY\_G\_RT1

Histogram



# CEM\_Min

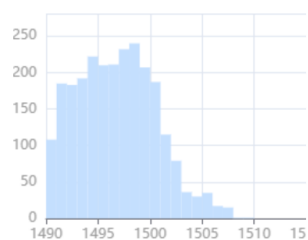
Histogram



Posterior Distributions

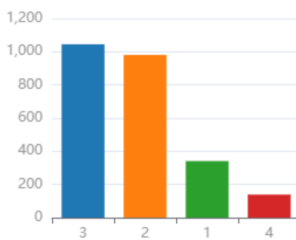
# WOC

Histogram



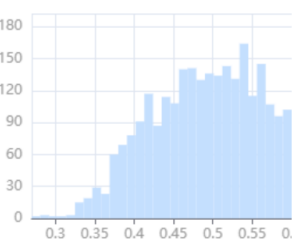
▲ FZI\_PHI

Histogram



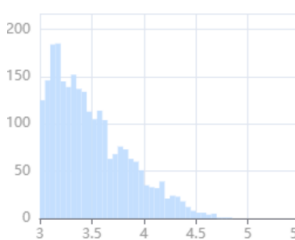
# KRW\_RT1

Histogram



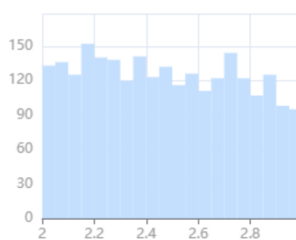
# COREY\_W\_RT1

Histogram



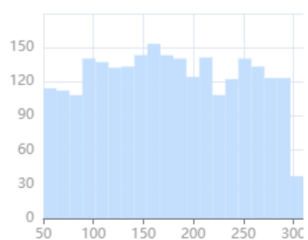
# COREY\_G\_RT1

Histogram

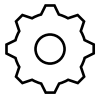


# CEM\_Min

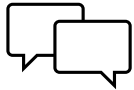
Histogram



# Co-development set up



Steering committee, governance team and project team are conformed by OMV and Schlumberger representatives



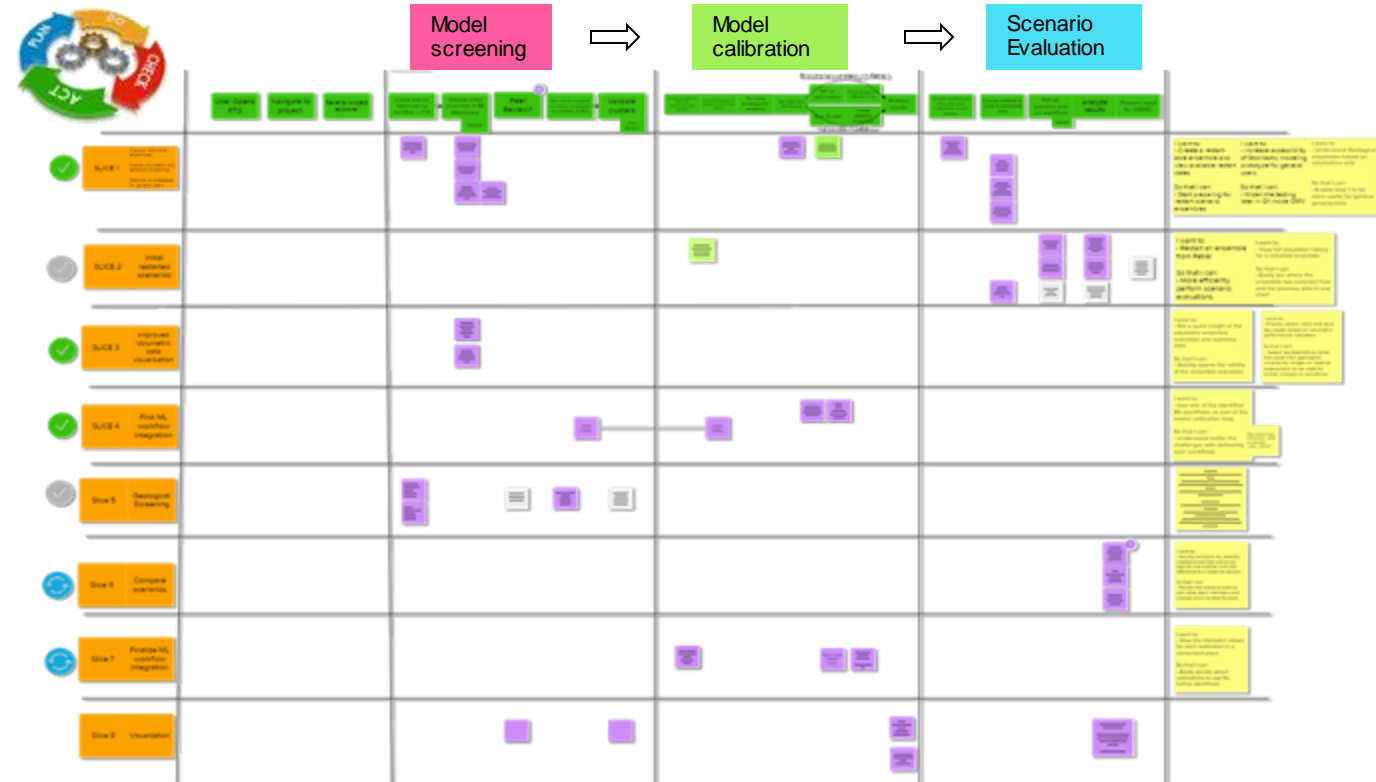
Bi-weekly sprint review meetings, every 6-month governance meetings and regular updates to the steering committee



Story map captures feedback and direction of development

Testing program based on OMV use cases

Constant dialogue with the development team



# The main benefits to OMV

---

OMV needs a stochastic modelling workflow to condition model ensembles to multiple data at well and reservoir level for **systematic and auditable** probabilistic production forecast profiles.

With the implementation of the Machine-Learning workflow in DELFI we have achieved:

- Reduction of working environments
- Simplification of data transferring
- Full integration of disciplines
- Parallelization of geomodelling and simulation
- Improved user experience from code to interactive interface
- Access to powerful visualization tools for data science
- Easier testing and implementation of alternative algorithms
- Flexible loop back mechanisms (optimization)
- Easier deployment to field business units





# Thank you!

OMV Exploration & Production

**Gisela Vanegas Cabas**  
Department Manager Integrated Services  
E&P – Value Center Development  
[gisela.vanegascabas@omv.com](mailto:gisela.vanegascabas@omv.com)

The energy for a better life.   
**OMV**