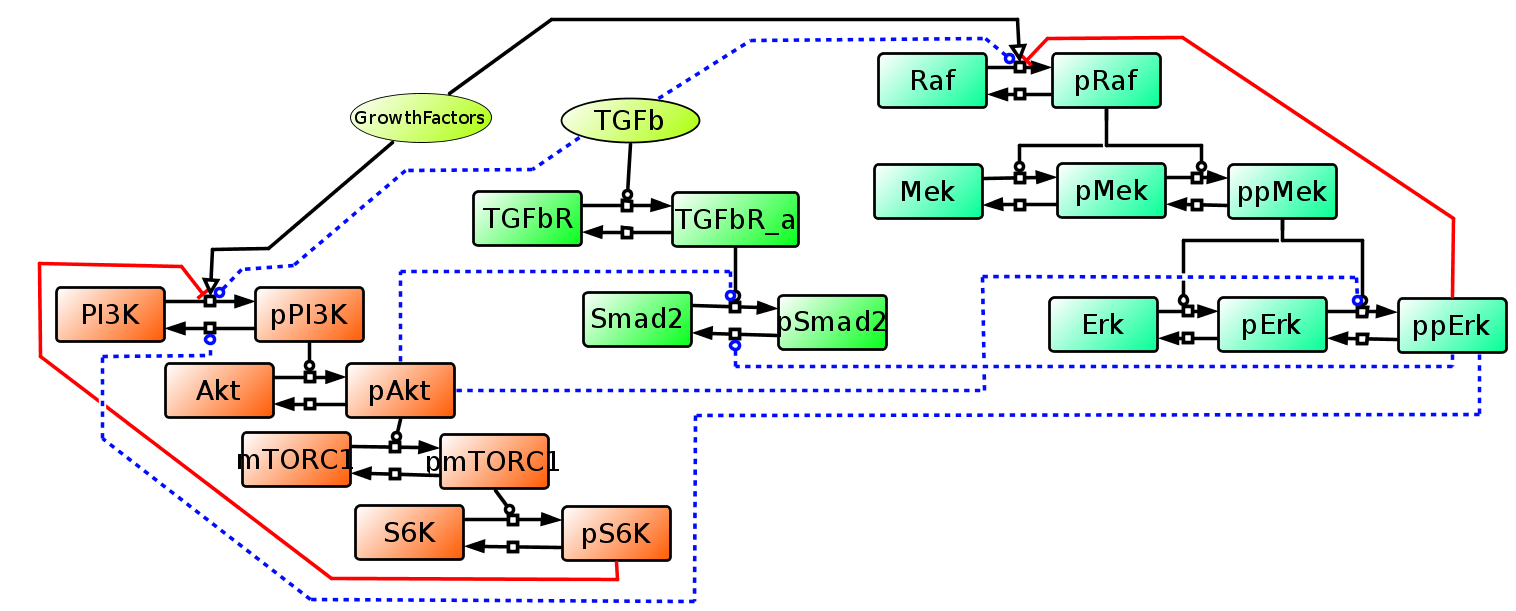
**2. Presentation of results**

The results are presented separately for computational models of signaling and metabolism.

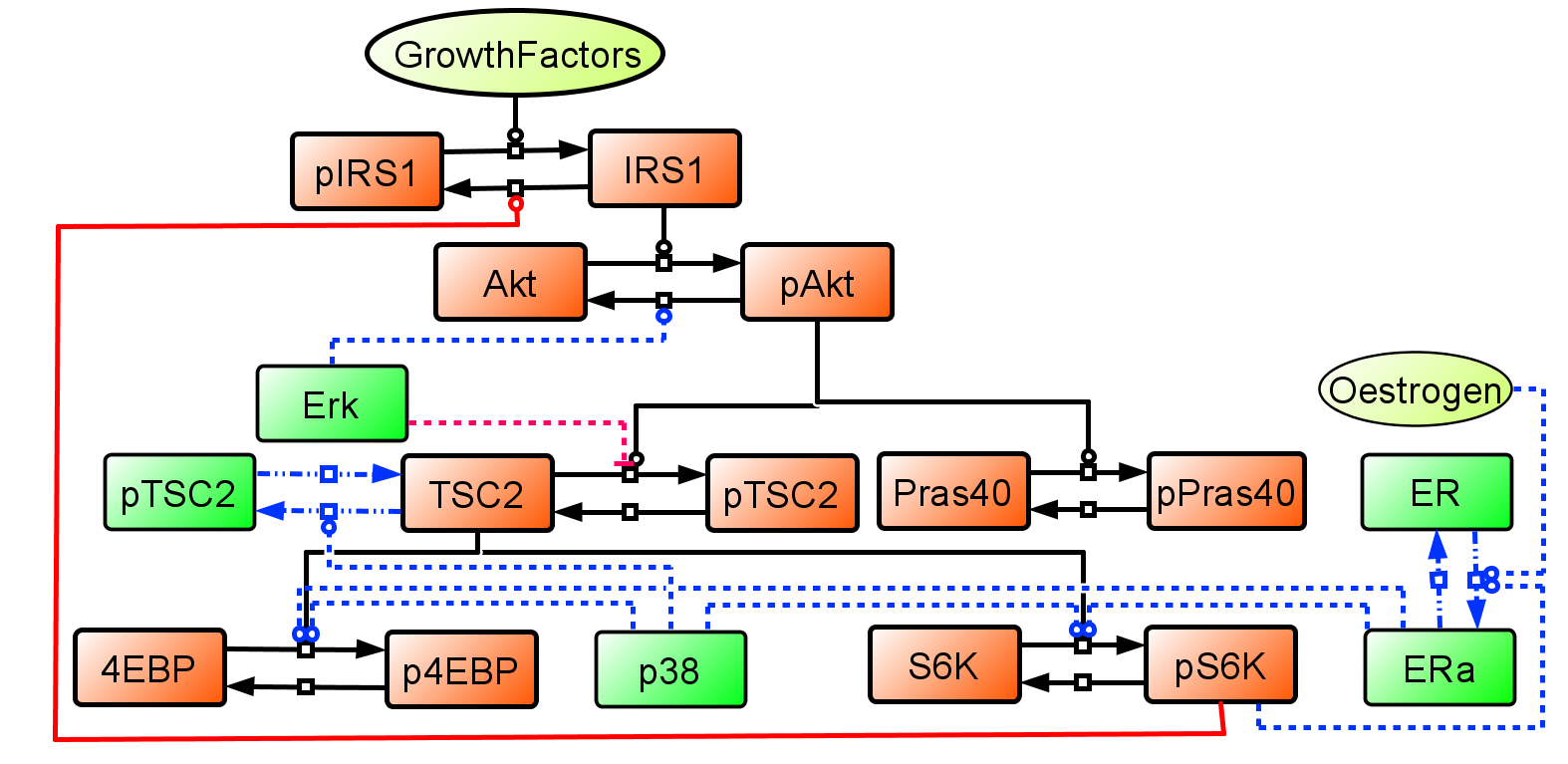
2.1 Signalling

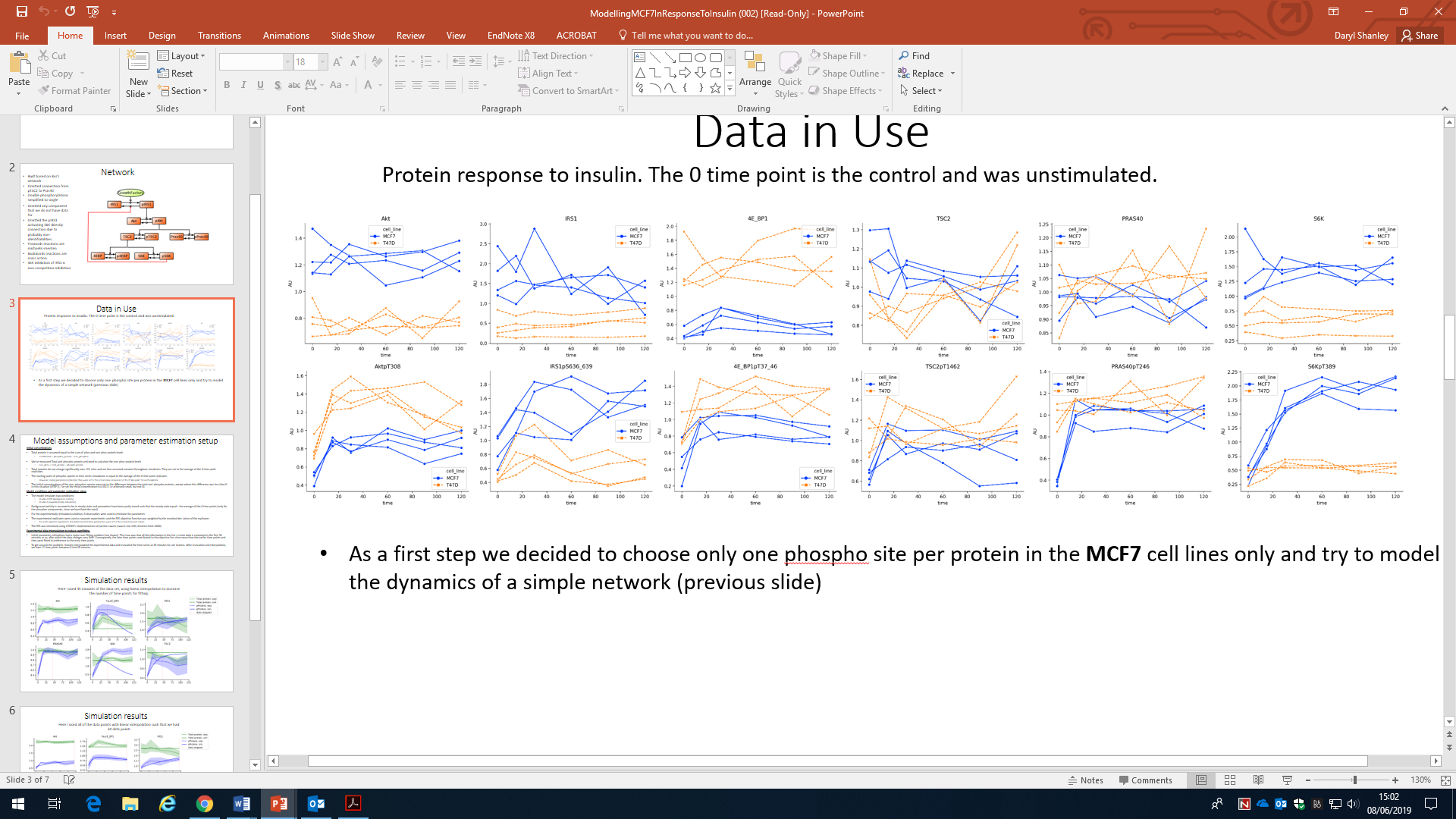


*Figure 1 Signalling network topology used to investigate crosstalk between PI3K/AKT/mTOR and MAPK networks and their input into TGFb signaling. Black lines show activation, red show inhibition, dashed blue lines show potential crosstalk between different parts of the network. These potential links are investigated singly and in combination by simulation using an automated modelling workflow.*

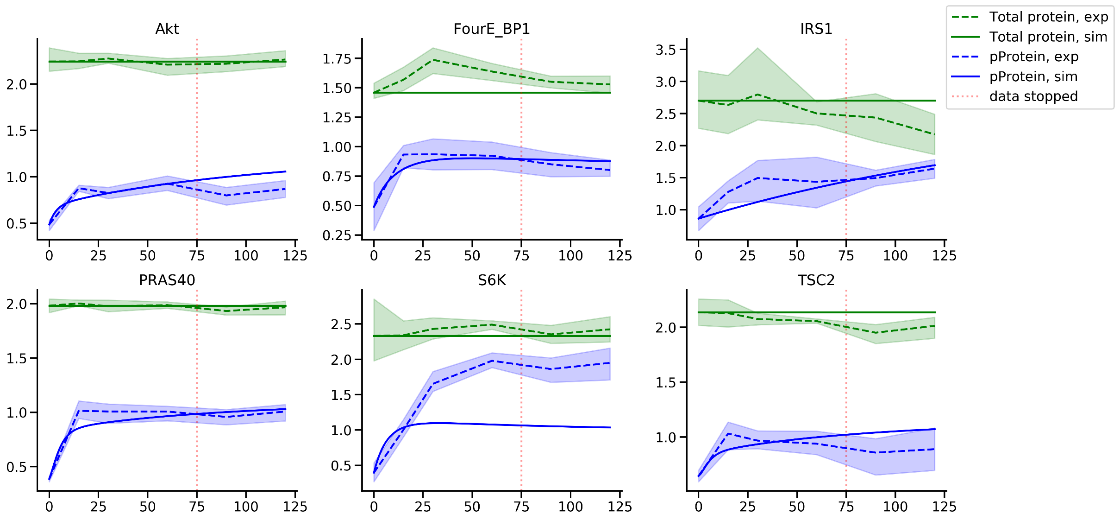
The integrated network shown in Figure 1 was developed for a related cancer biology project (MAPTor-NET) and represents one of several models used to inform the development of models for MESI-STRAT. Crosstalk between different parts of signaling networks is particularly difficult to disentangle through experiment alone and is one area where modelling has proved useful. The figure shows several places where crosstalk may be active and they are represented as alternative hypotheses within specific models. Model fitting is carried out for all combinations of hypotheses which for example in the case of 8 hypotheses results in 256 models. The fits are critically assessed and typically the top 5-10 best fitting models are taken forward for further investigation. For modelling fitting we use a several parameter estimation algorithms including global methods such as Particle Swarm and Simulated Annealing and local methods such as Hooke and Jeeves. Model fitting is supplemented with assessment of model identifiability and subsequent sensitivity analysis. This whole process is highly compute intensive and runs are automated on a HPC and interpretation of the output is managed within our PyCoTools software. Communication with FAIRDOM Hub using the RESTful API with appropriate ISA structures has been enbled within PyCoTools.

A



B**

C

**

*Figure 2 A ER-MAPK-mTOR signalling network topology. Black lines show activation, red show inhibition, dashed blue lines show potential crosstalk between different parts of the network. These potential links are investigated singly and in combination by simulation using an automated modelling workflow. B Experimental time course data for protein and phosphoprotein nodes in the signaling network measured in MCF7 and T47D cells. C Comparison of simulated and experimental data for selected nodes of the network. Note that for the results shown we only fit the model to the most informative dynamics which were in the first 75 minutes.*

**A COUPLE OF PARAGRAPHS ON WHAT YOU DID TO MODEL THE MCF7 DATA**.

To model the MCF7 cell line data we constructed a network representing the dynamics of the PI3K signalling pathway (Figure 2A, orange). The network was designed to be minimalistic, with a view to add the secondary phosphorylations once these were established. The network was designed to be stimulated by the insulin node, which was assumed to be constant on grounds that the signal is saturated over the time scales being simulated. The model simulates two conditions: 1) steady state, whereby insulin is equal to 0.005 (i.e. residual stimulation) and 2) dynamic, whereby insulin is set to 1 and kept constant.

The model was constructed using a combination of the antimony model definition language, COPASI, tellurium and PyCoTools. Michaelis-Menten rate laws were used for all forwards reactions while mass action was used for the backwards reactions. All kinetic parameters of the model were estimated from data by minimizing the residual sum of squares objective function, which quantifies the discrepancy between simulated and experimental data. Only the phospho-proteins were used as observables in parameter estimations because total proteins were assumed to be constant over the time period in question. This also allowed some simple algebra to be used to calculate the initial amount of non-phospho proteins (i.e. total = non-phosoho + phospho). To negate the possibility of phospho-proteins being present at higher levels than total proteins, an arbitrary value of 1 was added to the amounts of each non-phospho protein.

The average of the 0 time point measurements were used to calibrate the steady state condition, under the assumption that the cells were at steady state under this condition. For calibration of the model dynamics, each experimental repeat of each observable was used as independent experiments. This achieves two goals: 1) it enables the objective function to be weighted by the standard deviation of the experimental data and thus place more importance on low error data and 2) allows us to use the measured values at the 0 time points for each repeat of each observable as the initial concentration, thereby allowing use of measured values of initial concentration parameters.

The parameters were optimised using COPASI’s implementation of the particle swarm algorithm, configured using PyCoTools to increase efficiency and throughput. The simulations from the best parameter set (i.e. lowest RSS) are shown in Figure 2C.