# **Handover for Isoprene Research Assistantship**

Bikem Pastine, 31 May 2024

**Background:**

In this research assistantship, I continued the investigation into the viability of adapting the Energetic Status Model (Morfopolous et al., 2014) of isoprene emission for the global scale using the P-Model (Stocker et al., 2020). The initial investigation is described in my master's thesis titled “Modelling Global Isoprene Emissions with a Simplified Process-Based Approach Driven by the P-Model". During my master's research, I found that a barebones model where isoprene emissions are assumed to be a linear function of isoprene synthase activity and J (light limited electron flux), as calculated by the monthly version of the P-Model, is able to capture about 70% of the variation in satellite-based isoprene emission estimates (p<0.001). This investigation also revealed that the error of the simplified model (i.e. what was left unexplained) varied systematically with ambient temperature, shortwave radiation, and FPAR among other environmental variables, indicating that a full adaption of the Energetic Status Model could improve the explanatory power.

Given the relatively high proportion of the variation in isoprene emissions explained by the barebones model when compared to MEGAN (Guenther et al., 2006), a statistical and parameter dense model of isoprene, further investigation is warranted. This research assistantship focused on developing a method for adapting the full Energetic Status Model into a form that is compatible with the assumptions of the P-Model. The key challenge in compatibility between the two models is that the Energetic Status Model modulates emissions based on the balance between the total supply of photosynthetic reducing power (approximated as J) and the electron demand for carbon fixation (Jv). In the monthly version of the P-Model there is no difference in J and Jv due to the assumption that plants operate close to the intersection of light-limited and rubisco-limited assimilation rates (the coordination hypothesis). I addressed this incompatibility by approximating Jv with Jw – the theoretical electron flux generated in water stressed conditions.

Additionally, adapting the Energetic Status Model to the global scale poses a challenge in the calculation of the empirical scaling parameters c1 and c2. Because there are two scaling parameters, the dependencies and forms of the parameters are unknown, and there are additional constraints to the parameters (ie. c1 >> c2), it is necessary to develop a method for their calculation. A neural network-based approach was investigated during my research assistantship, but I concluded that this approach is ultimately not feasible whitin the scope of the research assistantship. The reasons for this conclusion are outlined below.

Finally, I investigated the ability of an updated, but still simplified model of isoprene emissions (aslo incorporating Jw) to reproduce emission extremes. This investigation is conducted considering the recent paper which finds that remotely sensed atmospheric formaldehyde can be used to detect climate extremes (Morfopolous et al., 2021). I find ................///

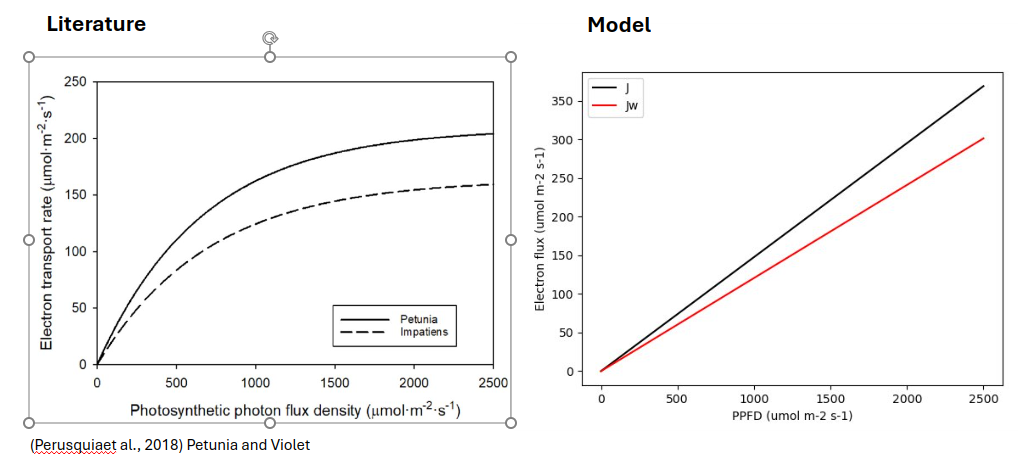
**Approximating Jv with Jw:**

The coordination hypothesis prevents the calculation of the energetic status as understood as [J – Jv]. However, we can assume that one of the main drivers of Jv falling significantly below J is water stress. This is because when there is water stress, stomata in the leaf close to reduce water loss by transpiration. The closure of stomata reduces the internal CO2 concentrations which thus limits carbon fixation by rubisco, meaning that the Calvin Cycle demands a lower electron flux while electron supply is unchanged.

The P-Model, like most vegetation productivity models, does not model the physiological mechanisms that reduce photosynthesis during water stress. Rather it applies a retrospective penalty (β) to its calculated GPP to reduce it by a given percentage. I assume that all the loss in photosynthesis approximated by β is a result of a reduction in the internal CO2 concentration. I calculate what the water stress internal CO2 concentration (WCi) must be at a given β to have the same GPP. I then use WCi to calculate Jw following the formula for Jv.

Where ci is the internal CO2 concentration, Vcmax is the maximum rate of carboxylation, Γ\* is the compensation point in the absence of dark respiration, kmm is the Michaelis-Menten constant for Rubisco-limited assimilation, and β is the water stress modifier from the SPLASH model.

While J and Jw, as calculated by the P-Model, captures the theoretical and experimental relationship between the energetic status and temperature as well as external CO2, it fails to correctly capture the shape of the relationship with PPFD and FPAR.



The reason for the mismatch is the optimality criterion in the P-Model. The P-Model assumes that on the time scale of acclimation (as in the monthly model version) Jmax adjusts to Iabs, so an increase in Iabs means that Jmax increases meaning that Aj and therefore J increase in line. We expect that there is some acclimation in isoprene production in the form of the shifting distribution of plants with the isoprene emission trait when there are prolonged environmental conditions that favor it. However, isoprene has a near immediate response to changes in light intensity and therefore, theoretically, the acclimation of Jmax should not govern isoprene emission estimates. The impact of this abstraction on isoprene modelling is that at low PPFD and FPAR, [J – Jw] changes in line with theory but at high light intensities [J – Jw] diverges from the known relationship.

In my opinion, the monthly version of the P-Model is not appropriate for isoprene emission modelling because of the short emission time of isoprene in plants. However, only monthly timestep global emission estimate data exists from satellites meaning that sub-daily isoprene estimates calculated using the P-Model will need to be averaged before being compared to observations, posing a statistical disadvantage. Ideally, fluxtower isoprene data could be used to validate the model at a smaller timestep before the global monthly model performance is assessed.

**Neural Network Exploration:**

The Energetic Status Model has the form:

Where I is isoprene flux, c1 and c2 are empirical scaling parameters representing the fraction of electrons directed towards isoprene synthesis, J is the light limited electron transport rate, and Jv is electron demand for carbon fixation.

The calculation of c1 and c2 is complex and ideally there would be an optimality-based approach to finding the controls on c1 and c2 so that a direct regression against observations would not be necessary. As a first pass towards identifying variables that may be of importance for determining the empirical parameters, I performed a lasso regression during my master's thesis. Based on this initial investigation it became apparent that some environmental variables seem to vary with c1 and that these relationships may not be straightforward.

The following goals were therefor defined:

* Identify environmental variables that drive the variation in c1 and c2.
* Infer the relationship between c1 and c2 and environmental variables in different conditions (i.e. different drivers may become more important in differing environmental contexts)
* Find the shape of relationships between environmental variables and the empirical parameters without assuming a linear relationship.

Because the shape of the relationships between c1 and c2 and environmental variables is unknown, flexible non-parametric estimation is preferred over a traditional regression model. A neural network is a popular non-parametric modelling tool and was considered as a candidate.

However, after some consideration and discussion with an expert (PhD candidate Jianing Fang) the conclusion emerged that this approach is not yet feasible. Unsolved issues with the NN approach include the overwhelming impact of selecting the environmental variables to investigate, difficulty in designing an investigation that would prioritize transparent interpretability over model fit using neural networks, and the current lack of knowledge about the distribution of the errors from both our observational data and the model.

**Extreme Detection:**

Using extremes in remotely sensed formaldehyde observations appears to be a promising method for identifying vegetation stress and climate extremes. Given that formadahyde is a byproduct of isoprene oxidation in the atmosphere, one can assume that isoprene could also be used to identify vegetative stress. For this reason, I investigated the ability of a simplified version of the Energetic Status Model to reproduce isoprene emission extremes detected in OMI data.

First, I decided on a form for the simplified version of the model to investigate. Below is a table of the R2 values calculated for the MEGAN model and three simplified model forms against OMI isoprene flux estimates.

|  |  |
| --- | --- |
| **Data Source** | **R2** |
| MEGAN and OMI | 0.75 |
| J × f(T) and OMI | 0.67 |
| J × f(T)× [J – Jw] and OMI | 0.07 |
| J×f(T) + J×f(T) × [J – Jw] and OMI | 0.53 |

P<0.001 in all cases. R2 calculated from the Pearson correlation coefficient.

I chose to investigate the extremes identified by the J×f(T) form as well as the J×f(T) + J×f(T) × [J-Jw] form because of their comparatively high R2 values.

The code and results I produced in this investigation have been passed forward allong with this document. I was unable to analyze the results.