****

**สถาบันวิจัยดาราศาสตร์แห่งชาติ**

**คู่มือการใช้งานโปรแกรม**

**เรื่อง Variable Star Template Fitting**

**โดย**

**นาย ประหยัด ปวงจักร์ทา**

**เสนอ**

**ดร. ฌหทัย ตนะกูล**

**สารบัญ**

[**1.** **โครงสร้างโปรเจค (Project Structure)** 3](#_Toc61317746)

[**1.1** **ส่วนประกอบของโปรแกรม** 4](#_Toc61317747)

[**2.** **การติดตั้ง (Install)** 7](#_Toc61317748)

[**2.1** **ขั้นตอนในการติดตั้ง (สำหรับ windows)** 7](#_Toc61317749)

[**3.** **การใช้งาน** 9](#_Toc61317750)

[**4.** **การปรับ Parameters** 12](#_Toc61317751)

# **โครงสร้างโปรเจค (Project Structure)**

ภาพด้านบนเป็นโครงสร้างของโปรเจคนี้ โดยประกอบด้วย 2 ส่วนหลัก คือส่วนของโปรแกรม (Program) และ ส่วนของเอกสาร (Doc) โดยในรูปภาพส่วนที่เป็นสี่เหลี่ยมจะแทนโฟลเดอร์และส่วนที่เป็นข้อมความจะแทนไฟล์ ซึ่งโฟลเดอร์จะถูกโยงด้วยลูกศร ส่วนไฟล์จะถูกโยงด้วยเส้นปะ ซึ่งส่วนประกอบของโปรเจคเหล่านี้แต่ละส่วนจะถูกอธิบายต่อไปดังนี้

## **ส่วนประกอบของโปรแกรม**

1. Doc Folder เป็นส่วนที่เก็บเอกสารต่างๆ สำหรับโปรเจค

2. Program Folder เป็นส่วนที่เก็บไฟล์ต่างๆ เกี่ยวกับโปรแกรมทั้งหมดโดยประกอบด้วยส่วนย่อยๆ หลายส่วนดังนี้

2.1 Data เป็นโฟล์เดอร์ที่เก็บข้อมูลนำเข้า (input) ของโปรแกรมโดยจะแบ่งเป็นสามโฟลเดอร์ย่อยได้แก่

- Data\_for\_fitlc เป็นโฟลเดอร์ที่เก็บข้อมูลนำเข้าสำหรับไฟล์ fitlc.py ซึ่งจะมีข้อมูลอยู่สามไฟล์ด้วยด้วยกัน ได้แก่

- ihjd.dat เก็บข้อมูลเวลาสำหรับ filter I

- vhjd.dat เก็บข้อมูลเวลาสำหรับ filter v

- montage.raw เก็บข้อมูล magnitude สำหรับ filter I, v ของแต่ละดวงดาว

- Data\_for\_rrfit เป็นโฟลเดอร์เก็บข้อมูลนำเข้าสำหรับไฟล์ template- fitting.py โดยประกอบด้วยสองส่วนย่อยได้แก่

- template คือโฟลเดอร์ที่เก็บ template ของดวงดาวแต่ละ แบบไว้ใช้สำหรับการ fitting ข้อมูล ด้านในประกอบด้วย ไฟล์ csv และ ไฟล์ txt ซึ่งทั้งสองไฟล์มีข้อมูลเหมือนกัน สำหรับโปรแกรมนี้ใช้ไฟล์ csv ในการ fitting เพราะความสะดวกในการจัดการกับข้อมูล

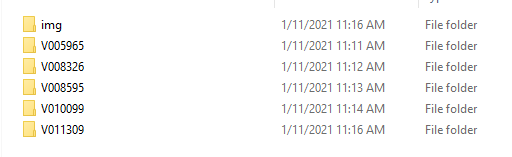
- fitlc\_star\_name\_txt.txt คือไฟล์รายชื่อดวงดาวทั้งหมดที่จะ ถูก fitting ในการ run แต่ละครั้ง โดยรายชื่อดวงดาวเหล่านี้จะผ่านการ กรองจากไฟล์ fitlc.py มาก่อนแล้ว

- Data\_from\_fitlc เป็นโฟลเดอร์ที่เก็บข้อมูลของดวงดาวแต่ละดวง ที่ ผ่านการกรองจากไฟล์ fitlc.py

2.2 lib เป็นโฟลเดอร์สำหรับเก็บ library ที่ได้เขียนขึ้นมาเอง ซึ่งผู้เขียนโปรแกรมใช้สำหรับการเก็บ function เพิ่มเติมที่จะนำไปใช้ไฟล์หลัก เพื่อลดความซับซ้อนของ main program และเพื่อความง่ายต่อการทำความเข้าใจในการพัฒนาต่อไป

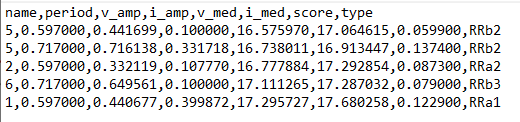
2.3 output คือโฟลเดอร์ที่เก็บผลลัพท์จากการ fitting ข้อมูลโดยแบ่งออกเป็นสองส่วนได้แก่

- img เป็นโฟลเดอร์ที่ถูกสร้างขึ้นมาเพื่อความสะดวกสบายแก่ผู้ตรวจความถูกต้อง ทำหน้าที่เก็บรูปภาพจากการ fitting data ของดวงดาวแต่ละดวงทั้งหมด

 - Star id โฟลเดอร์นี้เป็นชื่อสมมติเนื่องจากเวลารันโปรแกรมผลลัพท์จากการรันโปรแกรมนั้นจะทำการบันทึกข้อมูลแยกออกเป็นโฟลเดอร์ตามรหัสของดวงดาวดังรูป

2.4 requirements.txt เป็นไฟล์ text ที่เก็บข้อมูล library เสริมสำหรับโปรแกรมโดยจะต้องทำการติดตั้ง library เหล่านั้นก่อนทำการใช้โปรแกรม

2.5 all\_result.txt ไฟล์นี้เป็นไฟล์ text ที่เก็บข้อมูลผลลัพท์ที่ดีที่สุดจากการ fitting ของดวงดาวทุกดวงในการรันแต่ละครั้งไว้ ซึ่งข้อมูลแต่ละส่วนถูกคั่นด้วย comma ซึ่งสามารถนำไป import ในไฟล์ excel (ใช้การ delimit เพื่อแยก string) ได้ โดยมีตัวอย่างข้อมูลตามรูปภาพด้านล่าง

****

2.6 fitlc.py เป็นโปรแกรมที่ทำหน้าที่กรองดวงดาวที่มีความน่าจะเป็นจะเป็นดาวแปรแสง และจัดรูปแบบข้อมูลจากนั้นนำไปบันทึกไว้ในโฟลเดอร์ Data\_from\_fitlc เพื่อให้ไฟล์ change\_txt\_to\_csv.py และ template-fitting.py มาอ่านข้อมูลแล้วนำไปประมวลผลต่อไป

2.7 change\_txt\_to\_csv.py เป็นโปรแกรมที่ทำหน้าที่แปลงข้อมูลดวงดาวที่ผ่านการกรองจาก fitlc.py จากไฟล์ .txt ให้เป็น .csv (ซึ่งผลลัพท์ที่ได้จะเป็นไฟล์ text เพื่อความสะดวกในการอ่านข้อมูลและการจัดการข้อมูล ผู้จัดทำจึงต้องการแปลงข้อมูลให้อยู่ในรูปแบบไฟล์ csv)

2.8 temlate-fitting.py เป็นโปรแกรมหลักในการทำงาน โดยโปรแกรมจะทำงานโดยการหา period ที่เหมาะสมกับดวงดาวแต่ละดวงจากนั้นนำไป fitting เข้ากับ template ที่มีเพื่อตรวจสอบประสิทธิภาพแล้วทำการบันทึกค่าคลาดเคลื่อนนั้นไว้ เมื่อหาครบทุก period โปรแกรมจะเทียบประสิทธิภาพของการ fitting ในแต่ละ period ซึ่ง period ที่มีประสิทธิภาพมากที่สุดจะถูกเลือก แล้วนำไป plot เป็น graph แสดงผลการ fitting ในโฟลเดอร์ output และบันทึกผลการ fitting ต่างๆ ในโฟลเดอร์ output ด้วยเช่นกัน

# **การติดตั้ง (Install)**

โปรแกรมนี้ใช้ Library (โปรแกรมเสริมที่ผู้อื่นได้ทำการเขียนโปรแกรมเรียบร้อยแล้ว เพื่อช่วยอำนวยความสะดวกสบายในการเขียนโปรแกรมให้แก่ผู้อื่น) ที่ต้องทำการติดตั้งทั้งหมด 6 ตัว ได้แก่

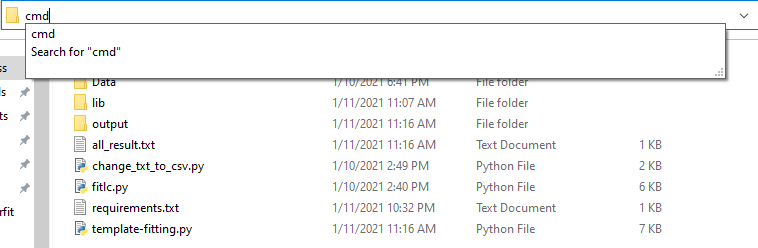
* Numpy
* Matplotlib
* Pandas
* Glob2
* Tqdm
* Scipy

ซึ่งทั้งหมดนี้ถูกบันทึกไว้ในไฟล์ที่มีชื่อว่า requirements.txt ซึ่งก่อนที่จะทำการใช้โปรแกรมต้องทำการติดตั้ง Library เหล่านี้ก่อน โปรแกรมจึงจะสามารถนำมาใช้ได้ โดยทำการติดตั้งผ่าน pip (คอมพิวเตอร์ที่จะนำโปรแกรมนี้ไปใช้งานต้องมีการติดตั้งภาษา python ก่อน เมื่อติดตั้ง python แล้วจะสามารถใช้ pip ได้อัติโนมัติ (หากไม่สำเร็จต้องทำการติดตั้ง python เข้าไปใน system path ก่อน))

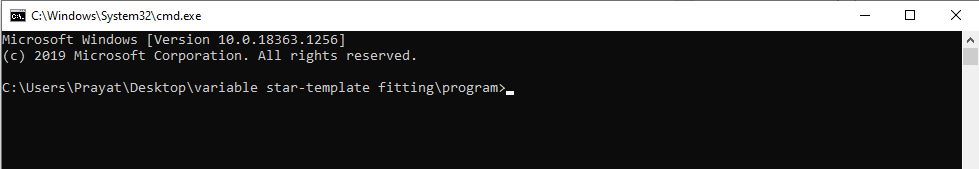
## **ขั้นตอนในการติดตั้ง (สำหรับ windows)**

1. เปิด Command Prompt

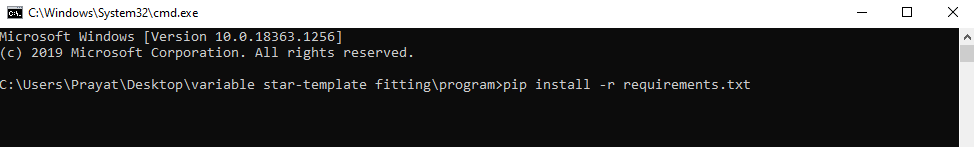
2. เข้าไปใน Folder ที่เป็นตัวโปรแกรม (ในที่นี้ผมใช้ชื่อ Folder ว่า “variable star-template fitting” ) สำหรับ windows อาจใช้การคลิกเข้าไปในโฟลเดอร์ “variable star-template fitting” จากนั้นคลิกต่อไปที่โฟลเดอร์ program แล้วพิมพ์ช่องด้านบนว่า “cmd” ตามรูป จากนั้นกด “Enter”



3.เมื่อเสร็จขั้นตอนที่ 2 จะมี Command Prompt ปรากฏมาโดยมีลักษณะดังรูปด้านล่าง



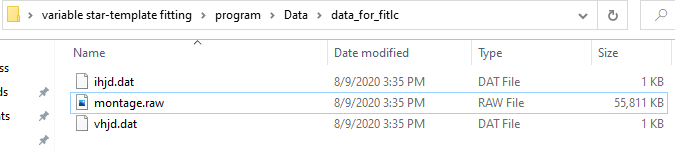
4. จากนั้นจะทำการพิมพ์ pip install -r requirements.txt จากนั้นทำการกด “Enter” โปรแกรมจะทำการดาวโหลด Library ทั้งหมดมาติดตั้งในคอมพิวเตอร์โดยอัติโนมัติ



กระบวนการทั้ง 4 ขั้นตอนหากทำทุกขั้นตอนเสร็จแล้ว ก็ถือว่าการติดตั้งเรียบร้อยสมบูรณ์ หลังจากนั้นก็จะสามารถใช้โปรแกรมที่เขียนมาได้ ต่อไปเป็นโครงสร้างของโปรเจค

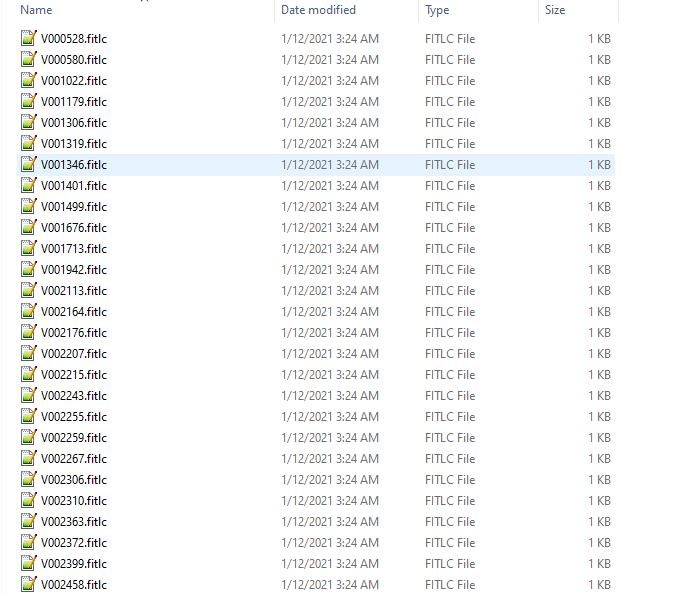
# **การใช้งาน**

1. การทำงานเริ่มต้นจากนำไฟล์ข้อมูลของดวงดาวที่ต้องการให้โปรแกรมรัน ไปวางไว้ที่โฟล์เดอร์ Data\_for\_fitlc

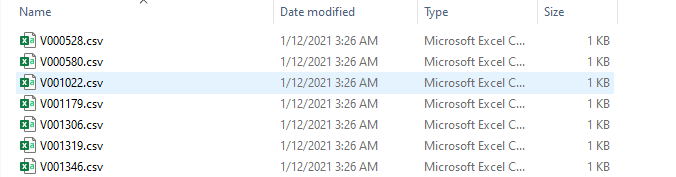


2. เมื่อนำไฟล์ข้อมูลนำเข้าไปวางไว้ใน Data\_for\_fitlc เสร็จเรียบร้อยแล้วให้ผู้ใช้โปรแกรมรันไฟล์ fitlc.py ไฟล์ fitlc.py จะทำการกรองข้อมูลดวงดาวขั้นต้นโดยใช้หลักการทางสถิติในการกรองและทำการจัดรูปแบบข้อมูลให้เหมาะสมสำหรับการ fitting จากนั้นบันทึกผลลัพท์ที่ได้ใน โฟล์เดอร์ Data\_from\_fitlc/txt

(การรันไฟล์ fitlc.py สามารถรันเหมือนกับไฟล์ python ปกติ)

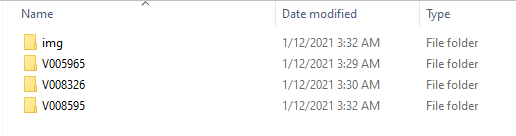
ผลลัพท์

3. จากนั้นให้การรันไฟล์ change\_txt\_to\_csv.py เพื่อทำการเปลี่ยนรูปแบบของไฟล์จาก text ให้กลายเป็นรูปแบบของ csv ผลลัพท์จะถูกบันทึกที่โฟล์เดอร์ Data\_from\_fitlc/csv ดังรูปด้านล่าง



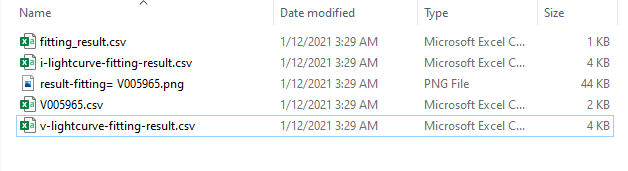
4. เมื่อขั้นตอนที่ 3 เสร็จสมบูรณ์ให้การรันไฟล์ template-fitting.py เมื่อเริ่มต้นรันจะมีไฟล์ all\_result.txt ปรากฎขึ้นมา และในส่วนของการแสดงผลจะมี progress bar แสดงจำนวนของดวงดาวที่ต้องรันทั้งหมดและบอกว่าตอนนี้ทำงานถึงดวงดาวดวงที่เท่าไหร่ดังรูปด้านล่าง



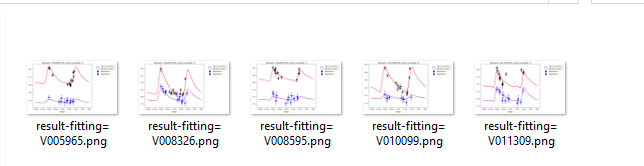
5. จากขั้นตอนที่ 4 จะใช้เวลาในการรันตามจำนวน period ที่ผู้ใช้โปรแกรมต้องการที่จะหา ซึ่งอาจจะใช้เวลาซักระยะหนึ่งในการทำงานในระหว่างที่ทำงานหากดาวดวงไหนที่ทำงานเสร็จแล้วสามารถเข้าไปตรวจสอบผลลัพท์ได้ที่โฟลเดอร์ output ซึ่งจะมีโฟล์เดอร์ของแต่ละดวงดาวปรากฎขึ้นมาในโฟล์เดอร์ดังรูปด้านล่าง

ด้านในของแต่ละโฟล์เดอร์จะประกอบด้วยไฟล์ผลลัพท์จากการ fitting อยู่ 5 ไฟล์ ได้แก่

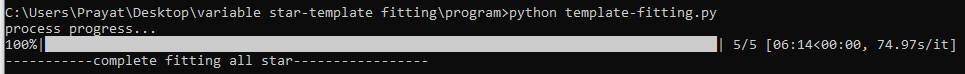
* fitting\_result.csv เป็นไฟล์ที่บอกผลลัพท์ของการ fitting ในแต่ละ template มีข้อมูลย่อยได้แก่ template, name, period, v\_amp, v\_median, i\_amp, i\_median, score (เป็นค่า error ยิ่งมีค่าน้อยแสดงว่าผลลัพท์ยิ่งดี)
* i-lightcurve-fitting-result.csv เป็นไฟล์ฟที่เก็บข้อมูลของ lightcurve filter i ที่ผ่านการ fitting กับ template เรียบร้อยแล้ว เป็นค่าข้อมูลที่แสดงในกราฟ
* v-lightcurve-fitting-result.csv เป็นไฟล์ฟที่เก็บข้อมูลของ lightcurve filter vที่ผ่านการ fitting กับ template เรียบร้อยแล้ว เป็นค่าข้อมูลที่แสดงในกราฟ
* result-fitting=star\_id.png เป็นรูปภาพของ graph ที่ plot ผลลัพท์ที่ดีที่สุดของการ fitting จาก lightcurve กับ magnitude(ที่ทำการปรับ period เรียบร้อย)
* V\_\_\_\_\_.csv เป็นไฟล์ที่เก็บข้อมูลทั้งหมดของดวงดาวและเพิ่ม period จาก template ที่ดีที่สุดของการ fitting



ส่วน output/img รูปจาก result-fitting=star\_id.png จะถูกนำไปบันทึกในโฟล์เดอร์นี้เพื่อความสะดวกในการตรวจสอบกราฟของแต่ละดวงดาว



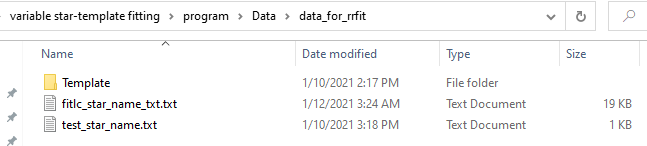
จากนั้นโปรแกรมจะนำผลลัพท์จากการ fitting บันทึกลงใน all\_result.txt สามารถตรวจสอบผลลัพท์ที่ดีที่สุดของแต่ละดวงดาวได้จากไฟล์นี้เมื่อโปรแกรมทำงานทั้งหมดเสร็จสิ้นจะมีการแสดงผลดังรูปด้านล่าง



# **การปรับ Parameters**

Parameters ที่สามารถปรับได้สำหรับไฟล์ template-fitting.py สำหรับการ fitting มีทั้งหมด 8 ตัว ได้แก่

* RUNNING\_FILE\_NAME คือไฟล์รายชื่อดวงดาวทั้งหมดที่จะให้โปรแกรม template-fitting.py รัน เปลี่ยนตามไฟล์รายชื่อดวงดาวทั้งหมดจาก Data/Data\_for\_rrfit/-------



โดย test\_star\_name.txt เป็นรายชื่อดวงดาวทั้งหมดสำหรับการ test หากใช้งานจริงให้แก้เป็น fitlc\_star\_name\_txt.txt

* ERAT เป็นตัวแปรที่มีหน้าที่เป็นค่า weight ของ magnitude error โดยโปรแกรมจะนำ ERAT ไปคูณกับ error เพื่อทำการปรับตามความเหมาะสมของแต่ละดวงดาว
* ERRMAX เป็นตัวแปรที่จำกัด magnitude error ของแต่ละดวงดาว หาก ณ ช่วงเวลาใดที่มี magnitude error เกินค่า ERRMAX โปรแกรมจะไม่นำข้อมูล ณ ช่วงเวลานั้นมาทำการคำนวณ
* EMIN เป็นตัวแปรที่กำหนดค่า magnitude error ต่ำสุด
* PERIOD1 เป็นค่าเริ่มต้นที่ใช้ในการหา period
* PERIOD2 เป็นค่าสิ้นสุดที่ใช้ในการหา period
* DPER เป็นค่า step ของการหา period เมื่อค่า DPER มาก ก็จะหา Period ได้น้อยใช้สำหรับเวลาต้องการหาผลลัพท์แบบคร่าวๆ ใช้เวลาในการ fitting น้อย แต่ DPER น้อย เป็นการหา Period ที่ละเอียดใช้เวลามากกว่า ผลลัพท์มีความละเอียดมากกว่า ตามสมการ

period = (1+dper)\*period1 ; หาจนกว่า period > period2

* TEMPLATE\_NUMBER เป็นตัวแปรสำหรับกำหนดจำนวน template ในการ fitting ข้อมูล



โดยปกติจะมี 8 template แต่หากต้องการลดลงให้ทำการเปลี่ยน -1 ให้น้อยลงไปเรื่อยๆ เช่น -2, -3, …. เป็นต้น

ภาพรวมของการปรับ parameters ในโปรแกรม โดยผู้จัดทำได้จัดให้การปรับค่าเหล่านี้ในไฟล์ template-fitting.py อยู่ส่วนบนของโปรแกรมผู้ใช้โปรแกรมสามารถค้นหา comment ที่มี Parameters กำกับอยู่ได้

